

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 12:49:27 ; Search time 127 seconds  
(without alignments)  
7163.856 Million cell updates/sec

Title: US-10-017-327-1

Perfect score: 1280

Sequence: 1 gaaagatggcgctccgcaag.....aaaaaaaaaaaaaaaaaaaa 1280

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	471.4	36.8	518	4	US-09-513-999C-12436 Sequence 12436, A
2	427.4	33.4	449	3	US-09-040-984-20 Sequence 20, Appl
3	427.4	33.4	449	3	US-09-123-912-20 Sequence 20, Appl
4	427.4	33.4	449	4	US-09-643-597-20 Sequence 20, Appl
5	427.4	33.4	449	4	US-09-480-884A-20 Sequence 20, Appl
6	427.4	33.4	449	4	US-09-542-615A-20 Sequence 20, Appl
7	427.4	33.4	449	4	US-09-606-421B-20 Sequence 20, Appl
8	427.4	33.4	449	4	US-09-221-107-20 Sequence 20, Appl
9	427.4	33.4	449	4	US-09-466-396A-20 Sequence 20, Appl
10	427.4	33.4	449	4	US-09-476-496A-20 Sequence 20, Appl
11	427.4	33.4	449	4	US-09-630-940B-20 Sequence 20, Appl
12	285.4	22.3	291	4	US-09-643-597-301 Sequence 301, Appl
13	285.4	22.3	291	4	US-09-480-884A-301 Sequence 301, Appl
14	285.4	22.3	291	4	US-09-542-615A-301 Sequence 301, Appl
15	285.4	22.3	291	4	US-09-606-421B-301 Sequence 301, Appl
16	285.4	22.3	291	4	US-09-630-940B-301 Sequence 301, Appl
17	236.4	18.5	2101	4	US-09-270-767-13882 Sequence 13882, A
18	89.4	7.0	457	4	US-09-513-999C-35509 Sequence 35509, A
19	72.8	5.7	5173	1	US-08-242-677-1 Sequence 1, Appl
20	72.6	5.7	2218	4	US-09-016-434-1157 Sequence 1157, Ap
21	72.6	5.7	2218	4	US-10-329-668-7 Sequence 7, Appl
22	72.6	5.7	2233	1	US-08-496-631-1 Sequence 1, Appl
23	72.4	5.7	396	4	US-09-640-173-53 Sequence 53, Appl
24	72.4	5.7	396	4	US-09-713-550-53 Sequence 53, Appl
25	72.4	5.7	396	4	US-09-825-294-53 Sequence 53, Appl
26	72.4	5.7	396	4	US-09-970-966-53 Sequence 53, Appl
27	71.8	5.6	1474	3	US-08-821-994-64 Sequence 64, Appl

28	71.4	5.6	1098	3	US-09-248-335-35 Sequence 35, Appl
29	71.4	5.6	1641	1	US-08-300-903A-8 Sequence 8, Appl
30	71.4	5.6	1641	4	US-08-988-197-8 Sequence 8, Appl
31	71.4	5.6	1641	4	US-10-385-072-8 Sequence 8, Appl
32	71.2	5.6	1492	4	US-09-369-247-23 Sequence 23, Appl
33	71.2	5.6	2671	6	5168051-9 Patent No. 5168051
34	71.2	5.6	4239	4	US-09-815-048-1 Sequence 1, Appl
35	70.6	5.5	396	4	US-09-640-173-16 Sequence 16, Appl
36	70.6	5.5	396	4	US-09-713-550-16 Sequence 16, Appl
37	70.6	5.5	396	4	US-09-825-294-16 Sequence 16, Appl
38	70.6	5.5	396	4	US-09-970-966-16 Sequence 16, Appl
39	70.6	5.5	1633	4	US-10-140-002-217 Sequence 217, App
40	70.6	5.5	4874	4	US-09-187-330-2 Sequence 2, Appl
41	70.4	5.5	997	4	US-09-907-794A-376 Sequence 376, App
42	70.4	5.5	997	4	US-09-905-125A-376 Sequence 376, App
43	70.4	5.5	997	4	US-09-902-775A-376 Sequence 376, App
44	70.4	5.5	997	4	US-09-906-700-376 Sequence 376, App
45	70.4	5.5	997	4	US-09-903-603A-376 Sequence 376, App

ALIGNMENTS

RESULT 1  
US-09-513-999C-12436  
; Sequence 12436, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59, US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 12436  
; LENGTH: 518  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: misc\_feature  
; NAME/KEY: misc\_feature  
; LOCATION: 3  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 92  
; OTHER INFORMATION: s=g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 493  
; OTHER INFORMATION: r=a or g  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 494  
; OTHER INFORMATION: r=a or g  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 495  
; OTHER INFORMATION: r=a or g  
; OTHER INFORMATION: 12436

Query Match 36.8%; Score 471.4; DB 4; Length 518;  
Best Local Similarity 98.8%; Pred. No. 7.4e-104;  
Matches 491; Conservative 4; Mismatches 0; Indels 2; Gaps 2;

Qy	1	GAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCGCTCTTCCAGCTCCACCG	60
Db	24	GAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCGCTCTTCCAGCTCCACCG	83

QY	61	CCGGCGCAGCAGGAAAGCAAGCAAGCGGCTCGGAGATTTCAGCCGTGAAGCAG	120
Db	84	CGCGCGCA-SAGGGAAAGCAAGCAAGCGGCTCGGAGATTTCAGCCGTGAAGCAG	142
QY	121	TCAGATAGATGGCTTGTGTTATTAAGATAATCAAAACATTATCAAGAAAGGACAAAG	180
Db	143	TCAGATAGATGGCTTGTGTTATTAAGATAATCAAAACATTATCAAGAAAGGACAAAG	202
QY	181	GAACTGAAAGTTGTTCAAGAGTGCTTTTGGTCTGGTTGTAGAGATCGGCTTGAATTA	240
Db	203	GAACTGAAAGTTGTTCAAGAGTGCTTTTGGTCTGGTTGTAGAGATCGGCTTGAATTA	262
QY	241	CCAACTGCTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT	300
Db	263	CCAACTGCTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT	322
QY	301	ATCAGATGGAATGATGCGGACCTTCCCATGTAACATTGATCATCTTTCAGGTGGCT	360
Db	323	ATCAGATGGAATGATGCGGA-CCTTCGCCATGTAACATTGATCATCTTTCAGGTGGCT	381
QY	361	GGTATCAGTCCACATACCTATGCTCATTCGTTACCGGCACTCTCGGACTCTCAGTTTA	420
Db	382	GGTATCAGTCCACATACCTATGCTCATTCGTTACCGGCACTCTCGGACTCTCAGTTTA	441
QY	421	GTTACCGATGCCATTGAAGAACTGTGCTTCTCATTTATGATCCCATATAAACTGCC	480
Db	442	GTTACCGATGCCATTGAAGAACTGTGCTTCTCATTTATGATCCCATATAARRCTGCC	501
QY	481	AAGGATCTCTCTCACTA	497
Db	502	AAGGATCTCTCTCACTA	518
RESULT 2			
US-09-040-984-20			
; Sequence 20, Application US/09040984			
; Patent No. 6210883			
; GENERAL INFORMATION:			
; APPLICANT: Reed, Steven G.			
; APPLICANT: Wang, TongTong			
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS			
; TITLE OF INVENTION: OF LUNG CANCER			
; NUMBER OF SEQUENCES: 86			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SEED AND BERRY LLP			
; STREET: 6300 Columbia Center, 701 Fifth Avenue			
; CITY: Seattle			
; STATE: WA			
; COUNTRY: USA			
; ZIP: 98104			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/040,984			
; FILING DATE: 18-MAR-1998			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Maki, David J.			
; REGISTRATION NUMBER: 31,392			
; REFERENCE/DOCKET NUMBER: 210121.456			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 206-622-4900			
; TELEFAX: 206-282-6031			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 20:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 449 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-09-040-984-20			
Query Match 33.4%; Score 427.4; DB 3; Length 449;			
Best Local Similarity 99.8%; Pred. No. 2.5e-93;			
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	804	ACTAGTAAACAACAGCAGCAGGAGAAACATCAGTATCAGCAGCGTCCGACAGGAGAAATATG	863
Db	1	ACTAGTAAACAACAGCAGCAGGAGAAACATCAGTATCAGCAGCGTCCGACAGGAGAAATATG	60
QY	864	CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA	923
Db	61	CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA	120
QY	924	CCACCACAGCCGCTGCCAGGATGGACTCGCTCATTTGCAGGCCAGATAAACACTTAC	983
Db	121	CCACCACAGCCGCTGCCAGGATGGACTCGCTCATTTGCAGGCCAGATAAACACTTAC	180
QY	984	TGCCAGAAATCAAGAGTTCACTGCCCAAAACTTAGGCAAGCTTTCATGGCCCAAGCT	1043
Db	181	TGCCAGAAATCAAGAGTTCACTGCCCAAAACTTAGGCAAGCTTTCATGGCCCAAGCT	240
QY	1044	CTTCAAGATAACAACAACCTAAGAAAGGAGTTTCCAGAAAAGAGTTTAAACATGAACCT	1103
Db	241	CTTCAAGATAACAACAACCTAAGAAAGGAGTTTCCAGAAAAGAGTTTAAACATGAACCT	300
QY	1104	TGAAGTCACACAGGCGCAACTCTTGGAAAGAAATATTTGCATATTTGAAAAGCAGAGG	1163
Db	301	TGAAGTCACACAGGCGCAACTCTTGGAAAGAAATATTTGCATATTTGAAAAGCAGAGG	360
QY	1164	ATTTCTTTAGTGTCAATGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAAT	1223
Db	361	ATTTCTTTAGTGTCAATGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAAT	420
QY	1224	AAAAAAAAA 1232	
Db	421	AAAAA 429	
RESULT 3			
US-09-123-912-20			
; Sequence 20, Application US/09123912A			
; Patent No. 6312695			
; GENERAL INFORMATION:			
; APPLICANT: Reed, Steven G.			
; APPLICANT: Wang, TongTong			
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER			
; FILE REFERENCE: 210121.455C1			
; CURRENT APPLICATION NUMBER: US/09/123,912A			
; CURRENT FILING DATE: 1998-07-27			
; PRIOR APPLICATION NUMBER: 09/040,802			
; PRIOR FILING DATE: 1998-03-18			
; NUMBER OF SEQ ID NOS: 114			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 20			
; LENGTH: 449			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-123-912-20			
Query Match 33.4%; Score 427.4; DB 3; Length 449;			
Best Local Similarity 99.8%; Pred. No. 2.5e-93;			
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	804	ACTAGTAAACAACAGCAGCAGGAGAAACATCAGTATCAGCAGCGTCCGACAGGAGAAATATG	863
Db	1	ACTAGTAAACAACAGCAGCAGGAGAAACATCAGTATCAGCAGCGTCCGACAGGAGAAATATG	60
QY	864	CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA	923
Db	61	CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA	120



Qy 924 CCACCACAGCGCCTGCCAGGATGGAATCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 983  
Db 121 CCACCACAGCGCCTGCCAGGATGGAATCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 180  
Qy 984 TGCCAGAACATCAAGGAGTTCACTGCCCAAACTTTAGCCAAGCTCTTCATGGCCCGAGCT 1043  
Db 181 TGCCAGAACATCAAGGAGTTCACTGCCCAAACTTTAGCCAAGCTCTTCATGGCCCGAGCT 240  
Qy 1044 CTTCAAGAATACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAAGTTAAACATGAACCTCT 1103  
Db 241 CTTCAAGATACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAAGTTAAACATGAACCTCT 300  
Qy 1104 TGAAGTCACACAGGCGAACTCTTGGGAAGAAATATATTTGTCATATTTGAAAAGCACAGAG 1163  
Db 301 TGAAGTCACACAGGCGAACTCTTGGGAAGAAATATATTTGTCATATTTGAAAAGCACAGAG 360  
Qy 1164 ATTTCTTTAGTGTCAATTTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATATAAAT 1223  
Db 361 ATTTCTTTAGTGTCAATTTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATATAAAT 420  
Qy 1224 AAAAAAAAA 1232  
Db 421 AAAAAAAAA 429

## RESULT 4

US-09-643-597-20  
; Sequence 20, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-643-597-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 804 ACTAGTAAACAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGATATG 863  
Db 1 ACTAGTAAACAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGATATG 60  
Qy 864 CAGCGCAGAGCCGAGGAGAACCCCGCTCCCTCAGGAGGACCTGTCCAACTCTTCAA 923  
Db 61 CAGCGCAGAGCCGAGGAGAACCCCGCTCCCTCAGGAGGACCTGTCCAACTCTTCAA 120  
Qy 924 CCACCACAGCGCCTGCCAGGATGGAATCGCTGCTCATTTGAGCCAGATAAACACTTAC 983  
Db 121 CCACCACAGCGCCTGCCAGGATGGAATCGCTGCTCATTTGAGCCAGATAAACACTTAC 180  
Qy 984 TGCCAGAACATCAAGGAGTTCACTGCCCAAACTTTAGCCAAGCTCTTCATGGCCCGAGCT 1043  
Db 181 TGCCAGAACATCAAGGAGTTCACTGCCCAAACTTTAGCCAAGCTCTTCATGGCCCGAGCT 240

Qy 1044 CTTCAAGAATACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAAGTTAAACATGAACCTCT 1103  
Db 241 CTTCAAGATACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAAGTTAAACATGAACCTCT 300  
Qy 1104 TGAAGTCACACAGGCGAACTCTTGGGAAGAAATATATTTGTCATATTTGAAAAGCACAGAG 1163  
Db 301 TGAAGTCACACAGGCGAACTCTTGGGAAGAAATATATTTGTCATATTTGAAAAGCACAGAG 360  
Qy 1164 ATTTCTTTAGTGTCAATTTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATATAAAT 1223  
Db 361 ATTTCTTTAGTGTCAATTTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATATAAAT 420  
Qy 1224 AAAAAAAAA 1232  
Db 421 AAAAAAAAA 429

## RESULT 5

US-09-480-884A-20  
; Sequence 20, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-480-884A-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 804 ACTAGTAAACAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGATATG 863  
Db 1 ACTAGTAAACAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGATATG 60  
Qy 864 CAGCGCAGAGCCGAGGAGAACCCCGCTCCCTCAGGAGGACCTGTCCAACTCTTCAA 923  
Db 61 CAGCGCAGAGCCGAGGAGAACCCCGCTCCCTCAGGAGGACCTGTCCAACTCTTCAA 120  
Qy 924 CCACCACAGCGCCTGCCAGGATGGAATCGCTGCTCATTTGAGCCAGATAAACACTTAC 983  
Db 121 CCACCACAGCGCCTGCCAGGATGGAATCGCTGCTCATTTGAGCCAGATAAACACTTAC 180  
Qy 984 TGCCAGAACATCAAGGAGTTCACTGCCCAAACTTTAGCCAAGCTCTTCATGGCCCGAGCT 1043  
Db 181 TGCCAGAACATCAAGGAGTTCACTGCCCAAACTTTAGCCAAGCTCTTCATGGCCCGAGCT 240  
Qy 1044 CTTCAAGAATACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAAGTTAAACATGAACCTCT 1103  
Db 241 CTTCAAGATACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAAGTTAAACATGAACCTCT 300  
Qy 1104 TGAAGTCACACAGGCGAACTCTTGGGAAGAAATATATTTGTCATATTTGAAAAGCACAGAG 1163  
Db 301 TGAAGTCACACAGGCGAACTCTTGGGAAGAAATATATTTGTCATATTTGAAAAGCACAGAG 360  
Qy 1164 ATTTCTTTAGTGTCAATTTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATATAAAT 1223  
Db 361 ATTTCTTTAGTGTCAATTTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATATAAAT 420

QY 1224 AAAAAAAAA 1232  
| | | | |  
Db 421 AAACAAAA 429

## RESULT 6

US-09-542-615A-20  
; Sequence 20, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hoeken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-542-615A-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 ACTAGTAAACACAGCAGCAGCAACATCAGTATCAGCAGCGTGCAGGAGGAGATATG 863  
| | | | |  
Db 1 ACTAGTAAACACAGCAGCAGCAACATCAGTATCAGCAGCGTGCAGGAGGAGATATG 60  
| | | | |  
QY 864 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGAGCTGTCCAAACTCTTCAA 923  
| | | | |  
Db 61 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGAGCTGTCCAAACTCTTCAA 120  
| | | | |  
QY 924 CCACACAGCGCGCTGCCAGATGGACTCGCTGCTATTCAGCGCCAGATATAACACTTAC 983  
| | | | |  
Db 121 CCACACAGCGCGCTGCCAGATGGACTCGCTGCTATTCAGCGCCAGATATAACACTTAC 180  
| | | | |  
QY 984 TGCAGAACATCAAGGAGTTCACTGCCCAAACTTAGGCAAGCTCTTCATGCGCCAGGCT 1043  
| | | | |  
Db 181 TGCAGAACATCAAGGAGTTCACTGCCCAAACTTAGGCAAGCTCTTCATGCGCCAGGCT 240  
| | | | |  
QY 1044 CTTCAAGAATACAACTAAGAAAAAGGAGTTTCCAGAAAAAGATTAAACATGAACCTCT 1103  
| | | | |  
Db 241 CTTCAAGAATACAACTAAGAAAAAGGAGTTTCCAGAAAAAGATTAAACATGAACCTCT 300  
| | | | |  
QY 1104 TGAAGTCACACAGCGCACTCTTGGAGAAATATATTTGCATATTGAAAAAGCAGAGG 1163  
| | | | |  
Db 301 TGAAGTCACACAGCGCACTCTTGGAGAAATATATTTGCATATTGAAAAAGCAGAGG 360  
| | | | |  
QY 1164 ATTTCTTTAGTGTCTATTCGCGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 1223  
| | | | |  
Db 361 ATTTCTTTAGTGTCTATTCGCGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 420  
| | | | |  
QY 1224 AAAAAAAAA 1232  
| | | | |  
Db 421 AAACAAAA 429

## RESULT 7

US-09-606-421B-20  
; Sequence 20, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hoeken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-606-421B-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 ACTAGTAAACACAGCAGCAGCAACATCAGTATCAGCAGCGTGCAGGAGGAGATATG 863  
| | | | |  
Db 1 ACTAGTAAACACAGCAGCAGCAACATCAGTATCAGCAGCGTGCAGGAGGAGATATG 60  
| | | | |  
QY 864 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGAGCTGTCCAAACTCTTCAA 923  
| | | | |  
Db 61 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGAGCTGTCCAAACTCTTCAA 120  
| | | | |  
QY 924 CCACACAGCGCGCTGCCAGATGGACTCGCTGCTATTCAGCGCCAGATATAACACTTAC 983  
| | | | |  
Db 121 CCACACAGCGCGCTGCCAGATGGACTCGCTGCTATTCAGCGCCAGATATAACACTTAC 180  
| | | | |  
QY 984 TGCAGAACATCAAGGAGTTCACTGCCCAAACTTAGGCAAGCTCTTCATGCGCCAGGCT 1043  
| | | | |  
Db 181 TGCAGAACATCAAGGAGTTCACTGCCCAAACTTAGGCAAGCTCTTCATGCGCCAGGCT 240  
| | | | |  
QY 1044 CTTCAAGAATACAACTAAGAAAAAGGAGTTTCCAGAAAAAGATTAAACATGAACCTCT 1103  
| | | | |  
Db 241 CTTCAAGAATACAACTAAGAAAAAGGAGTTTCCAGAAAAAGATTAAACATGAACCTCT 300  
| | | | |  
QY 1104 TGAAGTCACACAGCGCACTCTTGGAGAAATATATTTGCATATTGAAAAAGCAGAGG 1163  
| | | | |  
Db 301 TGAAGTCACACAGCGCACTCTTGGAGAAATATATTTGCATATTGAAAAAGCAGAGG 360  
| | | | |  
QY 1164 ATTTCTTTAGTGTCTATTCGCGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 1223  
| | | | |  
Db 361 ATTTCTTTAGTGTCTATTCGCGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 420  
| | | | |  
QY 1224 AAAAAAAAA 1232  
| | | | |  
Db 421 AAACAAAA 429

## RESULT 8

US-09-221-107-20  
; Sequence 20, Application US/09221107  
; Patent No. 6660838  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER  
; FILE REFERENCE: 210121.455C2  
; CURRENT APPLICATION NUMBER: US/09/221,107  
; CURRENT FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 161  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapiens

```
Db 361 ATTCTTTAGTGTCAATGGCGATTTTGGCTATACAGTGTCTTTAGCCATAATAAAT 420
QY 1224 AAAAAAAAA 1232
Db 421 AAAAAAAA 429

RESULT 11
US-09-630-940B-20
; Sequence 20, Application US/09630940B
; Patent No. 6737514
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C10
; CURRENT APPLICATION NUMBER: US/09/630,940B
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-630-940B-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;
Best Local Similarity 99.8%; Pred. No. 2.5e-93;
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 ACTAGTAACAACAGCAGCAGAGAAACATCAGTATCAGCAGCGTCGCAGCAGGAGAAATATG 863
Db 1 ACTAGTAACAACAGCAGCAGAGAAACATCAGTATCAGCAGCGTCGCAGCAGGAGAAATATG 60
QY 864 CAGCGCCAGCGCAGGAGAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA 923
Db 61 CAGCGCCAGCGCAGGAGAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA 120
QY 924 CCACCACAGCGCGCTGCCAGGATGGACTCGCTGCTCATTTGCGGCCAGATAAACACTTAC 983
Db 121 CCACCACAGCGCGCTGCCAGGATGGACTCGCTGCTCATTTGCGGCCAGATAAACACTTAC 180
QY 984 TGCAGAAATCAAGAGTTACTGCCCCAAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 1043
Db 181 TGCAGAAATCAAGAGTTACTGCCCCAAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 240
QY 1044 CTTCAAGATACAACTAAGAAAGGAGTTTCCAGAAACAGTAAACATGAACCTCT 1103
Db 241 CTTCAAGATACAACTAAGAAAGGAGTTTCCAGAAACAGTAAACATGAACCTCT 300
QY 1104 TGAAGTCACACGAGGCAACTCTTGAAGAAATATATTTGCATATTGAAAGCAGCAGAGG 1163
Db 301 TGAAGTCACACGAGGCAACTCTTGAAGAAATATATTTGCATATTGAAAGCAGCAGAGG 360
QY 1164 ATTCTTTAGTGTCAATGGCGATTTTGGCTATACAGTGTCTTTAGCCATAATAAAT 1223
Db 361 ATTCTTTAGTGTCAATGGCGATTTTGGCTATACAGTGTCTTTAGCCATAATAAAT 420

QY 1224 AAAAAAAAA 1232
Db 421 AAAAAAAA 429
```

```
RESULT 12
US-09-643-597-301/c
; Sequence 301, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 301
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-301

Query Match 22.3%; Score 285.4; DB 4; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.6e-59;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 946 TGGACTCGTGTCTATTGCGAGCCAGATAAACACTTACTGCCAGAACATCAAGGAGTTCA 1005
Db 291 TGGACTCGTGTCTATTGCGAGCCAGATAAACACTTACTGCCAGAACATCAAGGAGTTCA 232
QY 1006 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGGCCCCAGGCTCTTCAAGAATAACAACAACCTAG 1065
Db 231 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGGCCCCAGGCTCTTCAAGAATAACAACAACCTAG 172
QY 1066 AAAAGAAAGTTTCCAGAAAAGAGTTAAACATGAACCTCTTGAAGTCACACGAGGCAACTC 1125
Db 171 AAAAGAAAGTTTCCAGAAAAGAGTTAAACATGAACCTCTTGAAGTCACACGAGGCAACTC 112
QY 1126 TTGGAGAAATATATTGGCATATTGAAAGCAGAGGATTTCTTTAGTGTCTATTGCGCA 1185
Db 111 TTGGAGAAATATATTGGCATATTGAAAGCAGAGGATTTCTTTAGTGTCTATTGCGCA 52
QY 1186 TTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 1232
Db 51 TTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 5

RESULT 13
US-09-480-884A-301/c
; Sequence 301, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 301
```

US-09-221-107-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 ACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATG 863  
DB 1 ACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATG 60  
QY 864 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA 923  
DB 61 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA 120  
QY 924 CCACCAAGCGCCCTGCGAGGATGGACTCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 983  
DB 121 CCACCAAGCGCCCTGCGAGGATGGACTCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 180  
QY 984 TGCAGAAACATCAAGGAGTTCACCTGCGCCAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 1043  
DB 181 TGCAGAAACATCAAGGAGTTCACCTGCGCCAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 240  
QY 1044 CTTCAAGATACACACACACAACTAAGAAAGAAAGTTTCCAGAAAGAAAGTTTAACTGAACCTCT 1103  
DB 241 CTTCAAGATACACACACACAACTAAGAAAGAAAGTTTCCAGAAAGAAAGTTTAACTGAACCTCT 300  
QY 1104 TGAAGTCACACACAGGCAACTCTTGGAAAGAAATATATTTGCGATATTTGAAAGCAGAGG 1163  
DB 301 TGAAGTCACACACAGGCAACTCTTGGAAAGAAATATATTTGCGATATTTGAAAGCAGAGG 360  
QY 1164 ATTTCTTTAGTGTCAATGCGGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 1223  
DB 361 ATTTCTTTAGTGTCAATGCGGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 420  
QY 1224 AAAAAAAA 1232  
DB 421 AAAACAAA 429

RESULT 9

US-09-466-396A-20

; Sequence 20, Application US/09466396A  
; Patent No. 6696247  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C4  
; CURRENT APPLICATION NUMBER: US/09/466,396A  
; CURRENT FILING DATE: 1999-12-17  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-466-396A-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 ACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATG 863  
DB 1 ACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATG 60  
QY 864 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA 923  
DB 61 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGACCTGTCCAAACTCTTCAA 120  
QY 924 CCACCAAGCGCCCTGCGAGGATGGACTCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 983  
DB 121 CCACCAAGCGCCCTGCGAGGATGGACTCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 180

QY 984 TCCAGAAACATCAAGGAGTTCACCTGCCAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 1043  
DB 181 TCCAGAAACATCAAGGAGTTCACCTGCCAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 240  
QY 1044 CTTCAAGATACACACACAACTAAGAAAGAAAGTTTCCAGAAAGAAAGTTTAACTGAACCTCT 1103  
DB 241 CTTCAAGATACACACACAACTAAGAAAGAAAGTTTCCAGAAAGAAAGTTTAACTGAACCTCT 300  
QY 1104 TGAAGTCACACACAGGCAACTCTTGGAAAGAAATATATTTGCGATATTTGAAAGCAGAGG 1163  
DB 301 TGAAGTCACACACAGGCAACTCTTGGAAAGAAATATATTTGCGATATTTGAAAGCAGAGG 360  
QY 1164 ATTTCTTTAGTGTCAATGCGGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 1223  
DB 361 ATTTCTTTAGTGTCAATGCGGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 420  
QY 1224 AAAAAAAA 1232  
DB 421 AAAACAAA 429

RESULT 10

US-09-476-496A-20

; Sequence 20, Application US/09476496A  
; Patent No. 6706262  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF  
; TITLE OF INVENTION: LUNG CANCER  
; FILE REFERENCE: 210121.455C5  
; CURRENT APPLICATION NUMBER: US/09/476,496A  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 254  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-476-496A-20

Query Match 33.4%; Score 427.4; DB 4; Length 449;  
Best Local Similarity 99.8%; Pred. No. 2.5e-93;  
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 ACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATG 863  
DB 1 ACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATG 60  
QY 864 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCAA 923  
DB 61 CAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCAA 120  
QY 924 CCACCAAGCGCCCTGCGAGGATGGACTCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 983  
DB 121 CCACCAAGCGCCCTGCGAGGATGGACTCGCTGCTCATTTGCGAGCCAGATAAACACTTAC 180  
QY 984 TCCAGAAACATCAAGGAGTTCACCTGCCAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 1043  
DB 181 TCCAGAAACATCAAGGAGTTCACCTGCCAAACCTTAGGCAAGCTCTTCATGGCCCAAGGCT 240  
QY 1044 CTTCAAGATACACACACAACTAAGAAAGAAAGTTTCCAGAAAGAAAGTTTAACTGAACCTCT 1103  
DB 241 CTTCAAGATACACACACAACTAAGAAAGAAAGTTTCCAGAAAGAAAGTTTAACTGAACCTCT 300  
QY 1104 TGAAGTCACACACAGGCAACTCTTGGAAAGAAATATATTTGCGATATTTGAAAGCAGAGG 1163  
DB 301 TGAAGTCACACACAGGCAACTCTTGGAAAGAAATATATTTGCGATATTTGAAAGCAGAGG 360  
QY 1164 ATTTCTTTAGTGTCAATGCGGATTTGGCTATAACAGTGTCTTTAGCCATAATAAAT 1223

**This Page Blank (uspto)**

```
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-301

Query Match      22.3%; Score 285.4; DB 4; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.6e-59;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 946 TGGACTCGCTGCTCATTGCGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCA 1005
Db 291 TGGACTCGCTGCTCATTGCGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCA 232

QY 1006 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGAATAACAACAACCTAG 1065
Db 231 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGANTACAACAACCTAG 172

QY 1066 AAAGGAAGTTTCCAGAAAAGATTAAACATGAACCTCTTGAAGTCACACCGGGCAACTC 1125
Db 171 AAAGGAAGTTTCCAGAAAAGATTAAACATGAACCTCTTGAAGTCACACCGGGCAACTC 112

QY 1126 TTGGAAGAAATATATTGTCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGCCGA 1185
Db 111 TTGGAAGAAATATATTGTCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGCCGA 52

QY 1186 TTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 1232
Db 51 TTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 5

RESULT 14
US-09-542-615A-301/c
; Sequence 301, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Tasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 301
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-301

Query Match      22.3%; Score 285.4; DB 4; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.6e-59;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 946 TGGACTCGCTGCTCATTGCGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCA 1005
Db 291 TGGACTCGCTGCTCATTGCGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCA 232

QY 1006 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGAATAACAACAACCTAG 1065
Db 231 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGANTACAACAACCTAG 172

QY 1066 AAAGGAAGTTTCCAGAAAAGATTAAACATGAACCTCTTGAAGTCACACCGGGCAACTC 1125
Db 171 AAAGGAAGTTTCCAGAAAAGATTAAACATGAACCTCTTGAAGTCACACCGGGCAACTC 112

QY 1126 TTGGAAGAAATATATTGTCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGCCGA 1185
Db 111 TTGGAAGAAATATATTGTCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGCCGA 52

RESULT 15
US-09-606-421B-301/c
; Sequence 301, Application US/09606421B
; Patent No. 6511315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Tasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 301
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-301

Query Match      22.3%; Score 285.4; DB 4; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.6e-59;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 946 TGGACTCGCTGCTCATTGCGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCA 1005
Db 291 TGGACTCGCTGCTCATTGCGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCA 232

QY 1006 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGAATAACAACAACCTAG 1065
Db 231 CTGCCCCAAAACCTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGANTACAACAACCTAG 172

QY 1066 AAAGGAAGTTTCCAGAAAAGATTAAACATGAACCTCTTGAAGTCACACCGGGCAACTC 1125
Db 171 AAAGGAAGTTTCCAGAAAAGATTAAACATGAACCTCTTGAAGTCACACCGGGCAACTC 112

QY 1126 TTGGAAGAAATATATTGTCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGCCGA 1185
Db 111 TTGGAAGAAATATATTGTCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGCCGA 52

QY 1186 TTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 1232
Db 51 TTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 5

Search completed: January 5, 2005, 16:10:34
Job time : 129 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 13:00:26 ; Search time 746 Seconds  
(without alignments)

9691.377 Million cell updates/sec

Title: US-10-017-327-1

Perfect score: 1280

Sequence: 1 gaaagatggctccgcaag.....aaaaaaaaaaaaaaaa 1280

Scoring table: IDENTITY NUC

Gapop 10^0, Gapext 1.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1280	100.0	1280	10	US-09-878-328A-3
2	1280	100.0	1280	11	US-09-870-216C-1
3	1280	100.0	1280	13	US-10-017-327-1
4	1214.8	94.9	1337	9	US-09-925-300-413
5	1139.4	89.0	2285	10	US-09-814-353-20501
6	723.6	56.5	774	16	US-10-264-049-1286
7	619.4	48.4	621	9	US-09-777-564-1529
8	619.4	48.4	621	14	US-10-015-219-1529
9	563.4	44.0	565	11	US-09-969-034-3017
10	562.4	43.9	783	9	US-09-777-564-1299
11	562.4	43.9	783	14	US-10-015-219-1299
12	562	43.9	567	9	US-09-920-300A-168

c 13	562	43.9	567	13	US-10-033-528-168	Sequence 168, App
c 14	562	43.9	567	15	US-10-099-926-168	Sequence 168, App
c 15	514.4	40.2	571	11	US-09-969-034-1346	Sequence 1346, Ap
c 16	513.2	40.1	573	11	US-09-969-034-2625	Sequence 2625, Ap
c 17	501	39.1	723	16	US-10-428-681-23	Sequence 23, Appl
c 18	471.4	36.8	542	10	US-09-814-353-15325	Sequence 15325, A
c 19	454.4	35.5	469	16	US-10-242-535A-31398	Sequence 31398, A
c 20	454.4	35.5	469	16	US-10-085-783A-31398	Sequence 31398, A
c 21	451.6	35.3	651	11	US-09-969-034-4073	Sequence 4073, Ap
c 22	440.2	34.4	574	11	US-09-969-034-1482	Sequence 1482, Ap
c 23	438.8	34.3	597	11	US-09-969-034-1477	Sequence 1477, Ap
c 24	435	34.0	471	9	US-09-895-828-358	Sequence 358, App
c 25	435	34.0	471	15	US-10-114-866-358	Sequence 358, App
c 26	427.4	33.4	449	9	US-09-735-705-20	Sequence 20, Appl
c 27	427.4	33.4	449	9	US-09-850-116A-20	Sequence 20, Appl
c 28	427.4	33.4	449	9	US-09-897-778-20	Sequence 20, Appl
c 29	427.4	33.4	449	10	US-09-466-396A-20	Sequence 20, Appl
c 30	427.4	33.4	449	14	US-10-007-700-20	Sequence 20, Appl
c 31	427.4	33.4	449	15	US-10-117-982-20	Sequence 20, Appl
c 32	427.4	33.4	449	15	US-10-313-986-20	Sequence 20, Appl
c 33	427.4	33.4	449	18	US-10-775-972-20	Sequence 20, Appl
c 34	426.2	33.3	557	11	US-09-969-034-1461	Sequence 1461, Ap
c 35	412.8	32.2	454	9	US-09-880-107-2523	Sequence 2523, Ap
c 36	401	31.3	564	11	US-09-969-034-2642	Sequence 2642, Ap
c 37	387.4	30.3	466	10	US-09-918-995-29559	Sequence 29559, A
c 38	380.4	29.7	406	16	US-10-242-535A-48535	Sequence 48535, A
c 39	380.4	29.7	406	16	US-10-085-783A-48535	Sequence 48535, A
c 40	362	28.3	1130	16	US-10-264-049-533	Sequence 533, App
c 41	361.4	28.2	1149	14	US-10-097-065-116	Sequence 116, App
c 42	361.4	28.2	1149	15	US-10-372-876-116	Sequence 116, App
c 43	358.8	28.0	410	10	US-09-918-995-19562	Sequence 19562, A
c 44	348.8	27.3	410	9	US-09-960-352-4691	Sequence 4691, Ap
c 45	331.8	25.9	608	11	US-09-969-034-2836	Sequence 2836, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-878-328A-3

; Sequence 3, Application US/09878328A

; Publication No. US20030022174A1

; GENERAL INFORMATION:

; APPLICANT: VISAKORPI, TAPIO

; ISOLA, JORMA

; NUPPONEN, NINA

; OVOO, VOLODYMYR

; TITLE OF INVENTION: CANCER DIAGNOSTIC METHOD USING P40 SUBUNIT

; OF EIF3

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHVE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/878.328A

; FILING DATE: 12-Jun-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FI99/01039

; FILING DATE: 15-DEC-1999

; APPLICATION NUMBER: FI 982722

; FILING DATE: 16-DEC-1998

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

LENGTH: 1280 base pairs TYPE: nucleic acid STRANDEDNESS: single TOPOLOGY: linear SEQUENCE DESCRIPTION: SEQ ID NO: 3: US-09-878-328A-3									
Query Match Best Local Similarity 100.0%; Score 1280; DB 10; Length 1280; Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GAAGATGGCTCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG	60						
DB	1	GAAGATGGCTCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG	60						
QY	61	CCGCGCAGCAGGGAAGGCAAGGCAAGCGCGCTCGGAGATTCAGCGTGAAGCAAG	120						
DB	61	CCGCGCAGCAGGGAAGGCAAGGCAAGCGCGCTCGGAGATTCAGCGCTGAAGCAAG	120						
QY	121	TGCAGATAGATGGCTTGTGGTATTAAGATAATCAAAACATTATCAAGAAAGAGCAAG	180						
DB	121	TGCAGATAGATGGCTTGTGGTATTAAGATAATCAAAACATTATCAAGAAAGAGCAAG	180						
QY	181	GAACCTGAAGTTGTTCAAGGAGTGTCTTTGGGCTCTGGTTAGAAATTA	240						
DB	181	GAACCTGAAGTTGTTCAAGGAGTGTCTTTGGGCTCTGGTTAGAAATTA	240						
QY	241	CCAACTGCTTCCCTCCAGCAGCAGAGGATGCTGACTTTGATGAAGTCCAAT	300						
DB	241	CCAACTGCTTCCCTCCAGCAGCAGAGGATGCTGACTTTGATGAAGTCCAAT	300						
QY	301	ATCAGATGGAATGATCGGAGCTTCGCCATGTAACATTTGATCATCTTCACTGGGCT	360						
DB	301	ATCAGATGGAATGATCGGAGCTTCGCCATGTAACATTTGATCATCTTCACTGGGCT	360						
QY	361	GGTATCAGTCACATCTATGGCTCATCTGTTACCGGGCACTCTCGGACTCTCAGTTTA	420						
DB	361	GGTATCAGTCACATCTATGGCTCATCTGTTACCGGGCACTCTCGGACTCTCAGTTTA	420						
QY	421	GTTACCAAGATGCATTTGAAGATCTGCTCTCATTTATGATCCCATTAACACTGCC	480						
DB	421	GTTACCAAGATGCATTTGAAGATCTGCTCTCATTTATGATCCCATTAACACTGCC	480						
QY	481	AAGATCTCTCTCACTAAAGGCATACAGACTGACTCTAACTGATGGAAGTTTGAAG	540						
DB	481	AAGATCTCTCTCACTAAAGGCATACAGACTGACTCTAACTGATGGAAGTTTGAAG	540						
QY	541	AAAGGATTTTCCCTGAGCATTTGAAAAGCAATATACCTTTGAGTACATGTTTG	600						
DB	541	AAAGGATTTTCCCTGAGCATTTGAAAAGCAATATACCTTTGAGTACATGTTTG	600						
QY	601	AAGAAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGCTTAATGTGGAACTTG	660						
DB	601	AAGAAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGCTTAATGTGGAACTTG	660						
QY	661	AAAGAAGTCAGCTTTGCAGATAAATCAATGATGCTAGCTTGCAGCAATCAAT	720						
DB	661	AAAGAAGTCAGCTTTGCAGATAAATCAATGATGCTAGCTTGCAGCAATCAAT	720						
QY	721	TGGGGAAGATCTACAGTTGCTGATGGAAGAGTGGATGAATGAGCCAAAGATATAGTTA	780						
DB	721	TGGGGAAGATCTACAGTTGCTGATGGAAGAGTGGATGAATGAGCCAAAGATATAGTTA	780						
QY	781	AATACACATACATGAGGATCTAGTAAACACAGCAGCAGAAACATCAGTATCAGC	840						
DB	781	AATACACATACATGAGGATCTAGTAAACACAGCAGCAGAAACATCAGTATCAGC	840						
QY	841	AGCGTCGCCAGGAGGAATATGACGCGCAGAGCGGAGGAAACCCCGCTCCCTGAGG	900						
DB	841	AGCGTCGCCAGGAGGAATATGACGCGCAGAGCGGAGGAAACCCCGCTCCCTGAGG	900						
QY	901	AGGACTGTCCAAACTCTTCAACACACACAGCGCTGCCAGGATGACTCGCTGCTCA	960						
DB	901	AGGACTGTCCAAACTCTTCAACACACACAGCGCTGCCAGGATGACTCGCTGCTCA	960						

RESULT 2  
US-09-870-216C-1  
; Sequence 1, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1280  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-870-216C-1

Query Match Best Local Similarity 100.0%; Score 1280; DB 11; Length 1280; Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GAAGATGGCTCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG	60						
DB	1	GAAGATGGCTCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG	60						
QY	61	CCGCGCAGCAGGGAAGGCAAGGCAAGCGGCTCGGAGATTCAGCGTGAAGCAAG	120						
DB	61	CCGCGCAGCAGGGAAGGCAAGGCAAGCGGCTCGGAGATTCAGCGTGAAGCAAG	120						
QY	121	TGCAGATAGATGGCTTGTGGTATTAAGATAATCAAAACATTATCAAGAAAGAGCAAG	180						
DB	121	TGCAGATAGATGGCTTGTGGTATTAAGATAATCAAAACATTATCAAGAAAGAGCAAG	180						
QY	181	GAACCTGAAGTTGTTCAAGGAGTGTCTTTGGGCTCTGGTTAGAAATTA	240						
DB	181	GAACCTGAAGTTGTTCAAGGAGTGTCTTTGGGCTCTGGTTAGAAATTA	240						
QY	241	CCAACTGCTTCCCTCCAGCAGCAGAGGATGCTGACTTTGATGAAGTCCAAT	300						
DB	241	CCAACTGCTTCCCTCCAGCAGCAGAGGATGCTGACTTTGATGAAGTCCAAT	300						



241	DB	CCAACTGCTTTCCCTTCAGACACAGAGGATGATGCTGACCTTTGATGAAGTCCAA	300
301	QY	ATCAGATGGAATGATGCGGAGCCTTCGCCATGTAAACATTTGATCATCTTCACGTGGCT	360
301	DB	ATCAGATGGAATGATGCGGAGCCTTCGCCATGTAAACATTTGATCATCTTCACGTGGCT	360
361	QY	GGTATCAGTCACACATCTATGCTCATTTGTTTACCCGGGCACTCCTGGACCTCAGTTTA	420
361	DB	GGTATCAGTCACACATCTATGCTCATTTGTTTACCCGGGCACTCCTGGACCTCAGTTTA	420
421	QY	GTTTACCAGCATGCCATTTGAAGAACTGTGCTTCTCATTTATGATCCCAATAAAACCTGCC	480
421	DB	GTTTACCAGCATGCCATTTGAAGAACTGTGCTTCTCATTTATGATCCCAATAAAACCTGCC	480
481	QY	AAGGATCTCTCTCACTAAGGCATACAGACTGACTCTCTAAACTGATGGAAGTTGTAAAG	540
481	DB	AAGGATCTCTCTCACTAAGGCATACAGACTGACTCTCTAAACTGATGGAAGTTGTAAAG	540
541	QY	AAAAAGGATTTTTTCCCTGAAAGCATTTGAAAAAAGCAATATCACCTTTGAGTACATGTTTG	600
541	DB	AAAAAGGATTTTTTCCCTGAAAGCATTTGAAAAAAGCAATATCACCTTTGAGTACATGTTTG	600
601	QY	AAGAAGTGCAGATTGTAATTTAAAAATTCACATCTGATCAATGTCTCTAATGTGGGAACCTG	660
601	DB	AAGAAGTGCAGATTGTAATTTAAAAATTCACATCTGATCAATGTCTCTAATGTGGGAACCTG	660
661	QY	AAAAAGAGTCAGCTGTGTCAGATAACATGAATTTGCTCAGCGTTGCCAGCAGCAATCATT	720
661	DB	AAAAAGAGTCAGCTGTGTCAGATAACATGAATTTGCTCAGCGTTGCCAGCAGCAATCATT	720
721	QY	TGGGGAAGAACTTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCCAAGATATAGTTA	780
721	DB	TGGGGAAGAACTTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCCAAGATATAGTTA	780
781	QY	AATACAACACATACATGAGGAATCTAGTAAACAAACAGCAGCAACATCAGTATCAGC	840
781	DB	AATACAACACATACATGAGGAATCTAGTAAACAAACAGCAGCAACATCAGTATCAGC	840
841	QY	AGCGTCGCGCAGCAGGAATATGACGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAGG	900
841	DB	AGCGTCGCGCAGCAGGAATATGACGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAGG	900
901	QY	AGGACTGTCCAACTCTTCAAACCCACACAGCGCCTGCGCAGATGACTCGTGTCTCA	960
901	DB	AGGACTGTCCAACTCTTCAAACCCACACAGCGCCTGCGCAGATGACTCGTGTCTCA	960
961	QY	TTGCAGCGCAGATAAACATTTACTGCGCAGAACATCAAGGAGTTTCACTGCCCAAACTTAG	1020
961	DB	TTGCAGCGCAGATAAACATTTACTGCGCAGAACATCAAGGAGTTTCACTGCCCAAACTTAG	1020
1021	QY	GCAAGCTCTTCATGGCCAGGCTCTTCAAGAAATACAAACACTAAGAAAAAGGAAGTTTCCA	1080
1021	DB	GCAAGCTCTTCATGGCCAGGCTCTTCAAGAAATACAAACACTAAGAAAAAGGAAGTTTCCA	1080
1081	QY	GAAAAGAGTTTAAATGAATCTTTGAAGTCAACACAGGGCAACTCTTTGGAGAAATATAT	1140
1081	DB	GAAAAGAGTTTAAATGAATCTTTGAAGTCAACACAGGGCAACTCTTTGGAGAAATATAT	1140
1141	QY	TTGCATATTGAAAAGCAGAGGATTTCTTTAGTGTCACTTGCCCATTTTCGCTATAACAG	1200
1141	DB	TTGCATATTGAAAAGCAGAGGATTTCTTTAGTGTCACTTGCCCATTTTCGCTATAACAG	1200
1201	QY	TGCTCTTTCTAGCCATAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA	1260
1201	DB	TGCTCTTTCTAGCCATAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA	1260
1261	QY	AAAAAAAAAAAAAAAAAAAAA 1280	
1261	DB	AAAAAAAAAAAAAAAAAAAAA 1280	

```

; Sequence 1, Application US/10017327
; Publication No. US20020155471A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
; METHOD OF INVENTION: METHODS FOR USING SAME
; FILE REFERENCE: GZ 2101.20
; CURRENT APPLICATION NUMBER: US/10/017,327
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-327-1

```

Query Match 100.0%; Score 1280; DB 13; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 6.6e-288;  
Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GA	A	G	A	G	A	T	G	G	G	T	C	C	G	C	A	A	G	G	A	A	G	G	T	A	C	C	G	G	T	A	C	T	T	C	A	G	C	T	C	A	C	G	60
Db	1	GA	A	G	A	T	G	G	G	T	C	C	G	C	A	A	G	G	A	A	G	G	T	A	C	C	G	G	T	A	C	T	T	C	A	G	C	T	C	A	C	G	60		
Qy	61	CC	G	C	G	C	A	G	C	A	G	G	A	A	G	G	C	A	A	G	C	G	G	C	T	C	G	G	A	T	T	C	A	G	C	G	T	C	A	G	C	A	G	120	
Db	61	CC	G	C	G	C	A	G	C	A	G	G	A	A	G	C	A	A	G	C	G	G	C	T	C	G	G	A	T	T	C	A	G	C	G	T	C	A	G	C	A	G	120		
Qy	121	TG	C	A	G	A	T	A	G	C	T	T	G	T	A	A	A	G	A	T	A	A	A	C	A	T	A	T	A	C	A	A	G	A	A	G	A	C	A	A	G	180			
Db	121	TG	C	A	G	A	T	A	G	C	T	T	G	T	A	A	A	G	A	T	A	A	A	C	A	T	A	T	A	C	A	A	G	A	A	G	A	C	A	A	G	180			
Qy	181	GA	A	C	T	G	A	G	T	T	G	T	C	A	A	G	A	G	T	G	G	T	C	T	G	G	T	T	G	A	A	G	A	G	A	T	C	G	G	T	G	A	A	240	
Db	181	GA	A	C	T	G	A	G	T	T	G	C	T	C	A	A	G	A	G	T	G	G	T	C	T	G	G	T	T	G	A	A	G	A	T	C	G	G	T	G	A	A	240		
Qy	241	CC	A	A	C	T	G	C	T	T	C	C	C	T	C	A	G	C	A	C	A	G	A	G	A	T	G	C	T	G	A	A	G	A	T	G	C	A	A	T	G	300			
Db	241	CC	A	A	C	T	G	C	T	T	C	C	C	T	C	A	G	C	A	C	A	G	A	G	A	T	G	C	T	G	A	A	G	A	T	G	C	A	A	T	G	300			
Qy	301	A	T	C	A	G	A	T	G	A	T	C	C	G	A	G	C	C	T	C	G	C	A	T	G	T	A	A	A	C	A	T	T	G	A	T	C	A	T	C	T	C	360		
Db	301	A	T	C	A	G	A	T	G	A	T	C	C	G	A	G	C	C	T	C	G	C	A	T	G	T	A	A	A	C	A	T	T	G	A	T	C	A	T	C	T	C	360		
Qy	361	GG	T	A	T	C	A	G	T	C	C	A	T	A	T	G	G	T	C	A	T	T	G	T	T	A	C	C	G	G	A	C	T	C	T	G	A	C	T	C	T	420			
Db	361	GG	T	A	T	C	A	G	T	C	C	A	T	A	T	G	G	T	C	A	T	T	G	T	T	A	C	C	G	G	A	C	T	C	T	G	A	C	T	C	T	420			
Qy	421	GT	T	A	C	A	G	C	A	T	G	A	A	T	C	T	G	T	C	T	C	A	T	T	A	T	G	A	T	C	C	A	A	A	A	A	A	A	A	A	A	A	480		
Db	421	GT	T	A	C	A	G	C	A	T	G	A	A	T	C	T	G	T	C	T	C	A	T	T	A	T	G	A	T	C	C	A	A	A	A	A	A	A	A	A	A	A	480		
Qy	481	A	A	G	A	T	C	T	C	T	C	A	T	A	A	G	C	A	T	A	C	A	G	A	C	T	G	A	C	T	C	T	A	A	A	A	A	A	A	A	A	A	540		
Db	481	A	A	G	A	T	C	T	C	A	T	A	A	G	C	A	T	A	C	A	G	A	C	T	G	A	C	T	G	A															

QY 781 AATACACACATACATGAGGAATCTAGTAAACACAGCAGAGAAACATCAGTATCAGC 840  
DB |||||  
QY 781 AATACACACATACATGAGGAATCTAGTAAACACAGCAGAGAAACATCAGTATCAGC 840  
DB |||||  
QY 841 AGCGTGGCAGCAGGAGAAATGACGCGCAGAGCGAGAGAAACCCCGCTCCCTGAGG 900  
DB |||||  
QY 841 AGCGTGGCAGCAGGAGAAATGACGCGCAGAGCGAGAGAAACCCCGCTCCCTGAGG 900  
DB |||||  
QY 901 AGGACCTGTCCAAACTCTTCAACACACACAGCGCGCTCCGAGGATGACGTGCTCA 960  
DB |||||  
QY 901 AGGACCTGTCCAAACTCTTCAACACACACAGCGCGCTCCGAGGATGACGTGCTCA 960  
DB |||||  
QY 961 TTGACGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCACGTCCCAAACTTAG 1020  
DB |||||  
QY 961 TTGACGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCACGTCCCAAACTTAG 1020  
DB |||||  
QY 1021 GCAAGCTCTTCATGGCCGAGGCTCTTCAAGAAATCAACAACTAAGAAAGGAAGTTTCCA 1080  
DB |||||  
QY 1021 GCAAGCTCTTCATGGCCGAGGCTCTTCAAGAAATCAACAACTAAGAAAGGAAGTTTCCA 1080  
DB |||||  
QY 1081 GAAAGAGAGTTAAACATGAACTCTTGAAGTCAACAGCGGCACTCTTGGAGAAATATAT 1140  
DB |||||  
QY 1081 GAAAGAGAGTTAAACATGAACTCTTGAAGTCAACAGCGGCACTCTTGGAGAAATATAT 1140  
DB |||||  
QY 1141 TTGCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGGCCGATTTTGGCTATAACAG 1200  
DB |||||  
QY 1141 TTGCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGGCCGATTTTGGCTATAACAG 1200  
DB |||||  
QY 1201 TGTCTTTCTAGCCATATAAAATAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
DB |||||  
QY 1201 TGTCTTTCTAGCCATATAAAATAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
DB |||||  
QY 1261 AAAAAAAAAAAAAAAAAA 1280  
DB |||||  
QY 1261 AAAAAAAAAAAAAAAAAA 1280  
DB |||||

## RESULT 4

US-09-925-300-413  
; Sequence 413, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 413  
; LENGTH: 1337  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-300-413

Query Match 94.9%; Score 1214.8; DB 9; Length 1337;  
Best Local Similarity 98.2%; Pred. No. 1.1e-272;  
Matches 1247; Conservative 2; Mismatches 19; Indels 2; Gaps 2;  
QY 1 GAAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
DB |||||  
QY 34 GAAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 93  
DB |||||  
QY 61 CCGGCCAGCAGGAAAGCAAGCAAGCGGCTCGGAGATTCAGCCGTGGAAGCAAG 120  
DB |||||  
QY 94 CCGGCCCA-CAGGAAAGCAAGCAAGCGGCTCGGAGATTCAGCCGTGGAAGCAAG 152  
DB |||||  
QY 121 TGCAGATAGTGCCTTTGGTATTAAAGATTAATCAACATTAATCAAGAGAGGACAAAG 180  
DB |||||

DB 153 TGCAGATAGTGCCTTTGGTATTAAAGATTAATCAACATTAATCAAGAGAGGACAAAG 212  
QY |||||  
QY 181 GAACTGAAGTTGTTCAAGGAGTGTCTTTGGGTCTGGTTGTAGAGATCGGCTTGAATTA 240  
DB |||||  
QY 213 GAACTGAAGTTGTTCAAGGAGTGTCTTTGGGTCTGGTTGTAGAGATCGGCTTGAATTA 272  
DB |||||  
QY 241 CCAACTGCTTTCTCTTCCCTCAGCAGACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
DB |||||  
QY 273 CCAACTGCTTTCTCTTCCCTCAGCAGACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 332  
DB |||||  
QY 301 ATCAGATGGAATGATGCGGAGCTTCGCCATGTAAACATTAATGATCATCTTCAGTGGGCT 360  
DB |||||  
QY 333 ATCAGATGGAATGATGCGGA-SCITTCGCCATGTAAACATTAATGATCATCTTCAGTGGGCT 391  
DB |||||  
QY 361 GGTATCAGTCCACATACTATGGCTCATTCGTACCGGGCACTTCCTGGACTCTCAGTTTA 420  
DB |||||  
QY 392 GGTATCAGTCCACATACTATGGCTCATTCGTACCGGGCACTTCCTGGACTCTCAGTTTA 451  
DB |||||  
QY 421 GTTACAGCATGCGCATTTGAAGATCTGCTCTCATTTATGATCCCATATAAACTGCC 480  
DB |||||  
QY 452 GTTACAGCATGCGCATTTGAAGATCTGCTCTCATTTATGATCCCATATAAACTGCC 511  
DB |||||  
QY 481 AAGGATCTCTCCTCACTAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGAAG 540  
DB |||||  
QY 512 AAGGATCTCTCCTCACTAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGAAG 571  
DB |||||  
QY 541 AAGGATTTTTTCCCTGAAAGCATTTGAAAAAGCAAAATATACCTTTGAGTACATGTTTG 600  
DB |||||  
QY 572 AAGGATTTTTTCCCTGAAAGCATTTGAAAAAGCAAAATATACCTTTGAGTACATGTTTG 631  
DB |||||  
QY 601 AAGAGTGGCGATTGTAATTAATAATTCATCTGATCAATGTCTTAATGTGGGAACCTTG 660  
DB |||||  
QY 632 AAGAGTGGCGATTGTAATTAATAATTCATCTGATCAATGTCTTAATGTGGGAACCTTG 691  
DB |||||  
QY 661 AAGAGTGGCGATTGTAATTAATAATTCATCTGATCAATGTCTTAATGTGGGAACCTTG 720  
DB |||||  
QY 692 AAGAGTGGCGATTGTAATTAATAATTCATCTGATCAATGTCTTAATGTGGGAACCTTG 751  
DB |||||  
QY 721 TGGGGAAGATCTCAGTTGCTGATGAGCAGAGTGGATGAAATGAGCAGCAAGATATAGTTA 780  
DB |||||  
QY 752 TGGGGAAGATCTCAGTTGCTGATGAGCAGAGTGGATGAAATGAGCAGCAAGATATAGTTA 811  
DB |||||  
QY 781 AATACACACATACATGAGGAATCTAGTAAACACAGCAGAGAAACATCAGTATCAGC 840  
DB |||||  
QY 812 AATACACACATACATGAGGAATCTAGTAAACACAGCAGAGAAACATCAGTATCAGC 871  
DB |||||  
QY 841 AGCGTGGCAGCAGGAGAAATATGACGCGCAGAGCGAGAGAAACCCCGCTCCCTGAGG 900  
DB |||||  
QY 872 AGCGTGGCAGCAGGAGAAATATGACGCGCAGAGCGAGAGAAACCCCGCTCCCTGAGG 931  
DB |||||  
QY 901 AGGACCTGTCCAAACTCTTCAACACACAGCGCGCTTCGAGGATGAGCTCGCTGCTCA 960  
DB |||||  
QY 932 AGGACCTGTCCAAACTCTTCAACACACAGCGCGCTTCGAGGATGAGCTCGCTGCTCA 991  
DB |||||  
QY 961 TTGACGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCACGTCCCAAACTTAG 1020  
DB |||||  
QY 992 TTGACGGCCAGATAAACACTTACTGCCAGAACATCAAGAGTTCACGTCCCAAACTTAG 1051  
DB |||||  
QY 1021 GCAAGCTCTTCATGGCCGAGGCTCTTCAAGAAATCAACAACTAAGAAAGGAAGTTTCCA 1080  
DB |||||  
QY 1052 GCAAGCTCTTCATGGCCGAGGCTCTTCAAGAAATCAACAACTAAGAAAGGAAGTTTCCA 1111  
DB |||||  
QY 1081 GAAAGAGTTAAACATGAACTCTTGAAGTCAACAGCGGCACTCTTGGAGAAATATAT 1140  
DB |||||  
QY 1112 GAAAGAGTTAAACATGAACTCTTGAAGTCAACAGCGGCACTCTTGGAGAAATATAT 1171  
DB |||||  
QY 1141 TTGCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGGCCGATTTTGGCTATAACAG 1200  
DB |||||  
QY 1172 TTGCATATTGAAAGCAGAGGATTTCTTTAGTGTCAATGGCCGATTTTGGCTATAACAG 1231  
DB |||||  
QY 1201 TGTCTTTCTAGCCATATAAAATAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
DB |||||

Db 1232 TGCTCTTCTAGCCATAATAAAATAAAACAAAAATCTTGACTGCTGCTCAITTTTAAAAAAA 1291

Qy 1261 AAAAAAAAAA 1270

Db 1292 AAAAAAAAAA 1301

RESULT 5

US-09-814-353-20501

; Sequence 20501, Application US/09814353

; Publication No. US20030165831A1

; GENERAL INFORMATION:

; APPLICANT: Lee, John

; APPLICANT: Thompson, Pamela

; APPLICANT: Lillie, James

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353

; CURRENT FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-05-25

; PRIOR FILING DATE: 2000-06-15

; PRIOR FILING DATE: 2000-06-15

; PRIOR FILING DATE: 2000-07-07

; PRIOR FILING DATE: 2000-07-25

; PRIOR FILING DATE: 2000-07-25

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 22037

; SOFTWARE: PastSeq for Windows Version 4.0

; SEQ ID NO 20501

; LENGTH: 2285

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-814-353-20501

Query Match 89.0%; Score 1139.4; DB 10; Length 2285;

Best Local Similarity 92.7%; Pred. No. 5.3e-255;

Matches 1248; Conservative 0; Mismatches 26; Indels 73; Gaps 2;

Qy 1 GAAGATGGCTCC-CGCAAGGAGGTACCGGCTCTACTGCGACCTCTTCAGCTCCACC 59

Db 37 GAAAGATGGCTCCGCGCAAGGAGGTACCGGCTCTACTGCGACCTCTTCAGCTCCACC 96

Qy 60 GCCGCGCAGCAGGGAAGGCAAGGCAAGGCGGCTCGGAGATTGAGCGGTGAAGCAA 119

Db 97 GCCGCGCAGCAGGGAAGGCAAGGCAAGGCGGCTCGGAGATTGAGCGGTGAAGCAA 156

Qy 120 GTGAGATAGATGCGCTTGTGGTATTAAAGATAATCAAAATATCAAGGAAGGACAA 179

Db 157 GTGAGATAGATGCGCTTGTGGTATTAAAGATAATCAAAATATCAAGGAAGGACAA 216

Qy 180 GGAAGTGAAGTGTTCAGGAGTCTTTGGGCTGGTGTGAGAGATCGGCTTGAATTT 239

Db 217 GGAAGTGAAGTGTTCAGGAGTCTTTGGGCTGGTGTGAGAGATCGGCTTGAATTT 276

Qy 240 ACCAAGTCTTCTTTCCCTTCAGCAGAGATGATGCTGACTTTGATGAAG----- 294

Db 277 ACCAAGTCTTCTTTCCCTTCAGCAGAGATGATGCTGACTTTGATGAAGTTCA 336

Qy 295 ----- 294

Db 337 TTCATAGTCAGCATACATCAGAACTTCGTTCTTTGTTATTGCTGAGTAGTATTCATTA 396

Qy 295 -----TCCAAATCAGATGGAATGATGCGGAGCCTTCGCCATGTAAACATTTGATCAT 347

Db 397 AATGATTTCCAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAAACATTTGATCAT 456

RESULT 6

US-10-264-049-1286

; Sequence 1286, Application US/10264049

; Publication No. US20040005579A1

; GENERAL INFORMATION:

; APPLICANT: Birse et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA133P1

Qy 348 CTTCACTGGGCTGGTATCAGTCCACATATATATGGCTCATTTGTTTACCCGGGCACTCCTG 407

Db 457 CTTCACTGGGCTGGTATCAGTCCACATATATGGCTCATTTGTTTACCCGGGCACTCCTG 516

Qy 408 GACTCTCAGTTTATAGTTACCGAGCATGCCATTGAAGAACTCTGCTTCTCATTTATGATCCC 467

Db 517 GACTCTCAGTTTATAGTTACCGAGCATGCCATTGAAGAACTCTGCTTCTCATTTATGATCCC 576

Qy 468 ATAAAACTGCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATG 527

Db 577 ATAAAACTGCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATG 636

Qy 528 GAAGTTTGTAAAGAAAAGGATTTTCCCTGAAAGCATTTGAAAAAGCAAAATATCACCTTT 587

Db 637 GAAGTTTGTAAAGAAAAGGATTTTCCCTGAAAGCATTTGAAAAAGCAAAATATCACCTTT 696

Qy 588 GAGTACATGTTTGAAGAGTCCGATTGTAATTAATAAATTCACATCTGATCAATGTCTTA 647

Db 697 GAGTACATGTTTGAAGAGTCCGATTGTAATTAATAAATTCACATCTGATCAATGTCTTA 756

Qy 648 ATGTGGAACTTTGAAAAAGAGTCACTGCTGTCAGATAAAACATGAATTTGCTCAGCTTTGCC 707

Db 757 ATGTGGAACTTTGAAAAAGAGTCACTGCTGTCAGATAAAACATGAATTTGCTCAGCTTTGCC 816

Qy 708 AGCAGCAATCATTTGGGAAAGAACTCTACAGTTGCTGATGGACAGAGTGATGAAATGAGC 767

Db 817 AGCAGCAATCATTTGGGAAAGAACTCTACAGTTGCTGATGGACAGAGTGATGAAATGAGC 876

Qy 768 CAGATATAGTTTAATAACACATACATGATGAGGATCTAGTAAACACACAGCAGCAGAA 827

Db 877 CAGATATAGTTTAATAACACATACATGATGAGGATCTAGTAAACACACAGCAGCAGAA 936

Qy 828 CATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATGAGCGCCAGAGCCGAGGAGAACCC 887

Db 937 CATCAGTATCAGCAGCGTCGCCAGCAGGAGAAATATGAGCGCCAGAGCCGAGGAGAACCC 996

Qy 888 CCGCTCCCTGAGGAGGACCTGTGCAAACTCTTTCAAAACACACAGCGCCCTGCGAGGATG 947

Db 997 CCGCTCCCTGAGGAGGACCTGTGCAAACTCTTTCAAAACACACAGCGCCCTGCGAGGATG 1056

Qy 948 GACTCGCTGCTCATTTGAGGCGCAGATAAACACTTACTGCGCAGAACATCAAGGCTTCACT 1007

Db 1057 GACTCGCTGCTCATTTGAGGCGCAGATAAACACTTACTGCGCAGAACATCAAGGCTTCACT 1116

Qy 1008 GCCCAAACTTAGCAAGCTCTTCATGCGCCAGGCTCTTCAAGAAATCAACAACTAAGAA 1067

Db 1117 GCCCAAACTTAGCAAGCTCTTCATGCGCCAGGCTCTTCAAGAAATCAACAACTAAGAA 1176

Qy 1068 AAGGAAGTTTCCAGAAAAGGATTAACATGAACCTTTTGAAGTCAACACAGGCAACTCTT 1127

Db 1177 AAGGAAGTTTCCAGAAAAGGATTAACATGAACCTTTTGAAGTCAACACAGGCAACTCTT 1236

Qy 1128 GGAAGAAATATATTTGATATTTGAAAAGCAGAGGATTTCTTTAGTGTCTATTGCCGAT 1187

Db 1237 GGAAGAAATATATTTGATATTTGAAAAGCAGAGGATTTCTTTAGTGTCTATTGCCGAT 1296

Qy 1188 TTGCTTAACAGTGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAA 1247

Db 1297 TTGCTTAACAGTGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAA 1356

Qy 1248 AAAAAAAAAAAAAAAAAAAAAA 1274

Db 1357 CAGCAACAACACAGCACCACCAAAAA 1383

```

; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1286
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (686)..(686)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (722)..(722)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (769)..(769)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-1286

Query Match      56.5%; Score 723.6; DB 16; Length 774;
Best Local Similarity 97.4%; Pred. No. 2.5e-158;
Matches 755; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

QY 15 CGCAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGTCCACCGCGCGCAGCAGGG 74
Db 2 CGCAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGTCCACCGCGCGCAGCAGGG 61

QY 75 AAAGCAAGCAAGAGCGGCTCGGGAGATTCAGCCGTGAGCAAGTGCAGATAGATGCC 134
Db 62 AAAGCAAGCAAGAGCGGCTCGGGAGATTCAGCCGTGAGCAAGTGCAGATAGATGCC 121

QY 135 CTGTGGTATTAAAGATAATCAACATTATCAAGAGAGAGCAAGGAAGTGAAGTTGTT 194
Db 122 CTGTGGTATTAAAGATAATCAACATTATCAAGAGAGAGCAAGGAAGTGAAGTTGTT 181

QY 195 CAAGAGTGCTTTTGGGCTGGTGTAGAGATCGGCTTGAAGTACCAACTGCTTTTCT 254
Db 182 CAAGAGTGCTTTTGGGCTGGTGTAGAGATCGGCTTGAAGTACCAACTGCTTTTCT 241

QY 255 TTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAATATCAGATGGAATG 314
Db 242 TTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAATATCAGATGGAATG 301

QY 315 ATCGGAGCCTTCGCCATGTAACATGATCATCTTCACTGGGCTGGTATCAGTCCACA 374
Db 302 ATCGGAGCCTTCGCCATGTAACATGATCATCTTCACTGGGCTGGTATCAGTCCACA 361

QY 375 TACTATGGCTCATTCGTTTACCGGGCACTCTCGGACTCTCAGTTTGTAGTACGATGCC 434
Db 362 TACTATGGCTCATTCGTTTACCGGGCACTCTCGGACTCTCAGTTTGTAGTACGATGCC 421

QY 435 ATTGAAGAACTCTGCTGTTCTATTATGATGCCATAAAAGTCCCAAGGATCTCTCTCA 494
Db 422 ATTGAAGAACTCTGCTGTTCTATTATGATGCCATAAAAGTCCCAAGGATCTCTCTCA 481

QY 495 CTAAGGCATACAGACTGACTCTCAACTGATGGAAGTTTGTAAAGAAAGGATTTTTC 554
Db 482 CTAAGGCATACAGACTGACTCTCAACTGATGGAAGTTTGTAAAGAAAGGATTTTTC 541

QY 555 CTGAAGCAATTGAAAAAGCAATATCACTTTGATGATACATGTTTGAAGAAAGT-CCGAT 613
Db 542 CTGAAGCAATTGAAAAAGCAATATCACTTTGATGATACATGTTTGAAGAAAGTCCCGAT 601

QY 614 TGTAAATTAAATTCACATCTGATCAATGCTCTTAATGTGGGAAGTCTGAAAAGAGTCAGC 673
Db 602 TGTAAATTAAATTCACATCTGATCAATGCTCTTAATGTGGGAAGTCTGAAAAGAGTCAGC 661

```

```

QY 674 TGTTGCAGATAAACAATGAATTTGCTCAGCCTTGCAGCAGCAATCATTTGGGGAAGATCT 733
Db 662 TGTTGCAGATAAACAATGAATTTGCTCAGCCTTGCAGCAGCAATCATTTGGGGAAGATCT 719

QY 734 ACAGTTGCTGATGACAGAGTGCATGAATGAGCCAGATATATAGTTAAATATCAAC 788
Db 720 ACNGTTGCTGATGACCAAAATGATTAATAGCCAGGAATAGTAATNCCACC 774

RESULT 7
US-09-777-564-1529/c
; Sequence 1529, Application US/09777564
; Patent No. US2002022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Window Version 4.0
; SEQ ID NO 1529
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-777-564-1529

```

```

Query Match      48.4%; Score 619.4; DB 9; Length 621;
Best Local Similarity 99.8%; Pred. No. 4.4e-134;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 GCTCCACCGCGCGCGCAGCAGGAAAGCAAGGCGGCTCGGAGATTCAGCGG 111
Db 621 GCTCCACCGCGCGCGCAGCAGGAAAGCAAGGCGGCTCGGAGATTCAGCGG 562

QY 112 TGAAGCAAGTGCGATAGATGCGCTTGTGGTATTAAAGATAATCAACATTATCAAGAG 171
Db 561 TGAAGCAAGTGCGATAGATGCGCTTGTGGTATTAAAGATAATCAACATTATCAAGAG 502

QY 172 AAGCAAGGAACTGAAGTCTTCAAGAGAGTCTTTGGGCTCGTTGTAGAAAGATCGG 231
Db 501 AAGCAAGGAACTGAAGTCTTCAAGAGAGTCTTTGGGCTCGTTGTAGAAAGATCGG 442

QY 232 TTGAAATTACCACTGCTTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 291
Db 441 TTGAAATTACCACTGCTTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 382

QY 292 AAGTCCCAATATCAGATGGAATGATCGGAGCCTTCGCCATGTAAACATTGATCATCTTC 351
Db 381 AAGTCCCAATATCAGATGGAATGATCGGAGCCTTCGCCATGTAAACATTGATCATCTTC 322

QY 352 ACGTGGGCTGGTATCAGTCCACATFATGCTCATTCGTTACCCGGGCACTCTCTGACT 411
Db 321 ACGTGGGCTGGTATCAGTCCACATFATGCTCATTCGTTACCCGGGCACTCTCTGACT 262

QY 412 CTCAGTTTAGTTACCAGCATGCCATTGAAGAACTCTGCTTCATTTATGATCCCAATA 471
Db 261 CTCAGTTTAGTTACCAGCATGCCATTGAAGAACTCTGCTTCATTTATGATCCCAATA 202

QY 472 AAACGTCCCAAGGATCTCTCACTAAAGGCATACAGACTGACTCTTAACTGATGGAAG 531
Db 201 AAACGTCCCAAGGATCTCTCACTAAAGGCATACAGACTGACTCTTAACTGATGGAAG 142

QY 532 TTTGTAAGAAAAAGGATTTTCCCTCGAGCATTTGAAAAAGCAAAATATCACCTTTGAGT 591
Db 141 TTTGTAAGAAAAAGGATTTTCCCTCGAGCATTTGAAAAAGCAAAATATCACCTTTGAGT 82

QY 592 ACATGTTTGAAGAGTGGGATTTGTAATTAATAAATTCACATCTGATCAATGCTTAATGT 651
Db 81 ACATGTTTGAAGAGTGGGATTTGTAATTAATAAATTCACATCTGATCAATGCTTAATGT 22

```



```
RESULT 10
US-09-777-564-1299
; Sequence 1299, Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777.564
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Window Version 4.0
; SEQ ID NO 1299
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-777-564-1299

Query Match      43.9%; Score 562.4; DB 9; Length 783;
Best Local Similarity 97.6%; Pred. No. 9.7e-121;
Matches 603; Conservative 0; Mismatches 11; Indels 4; Gaps 3;

QY 52 GCTCCACCGCGCGCGAGCAAGGCAAGCGCGCTCGGAGATTCAGCG 111
Db 1 GCTCCACCGCGCGCGAGCAAGGCAAGCGCGCTCGGAGATTCAGCG 60

QY 112 TGAAGCAAGTGCAGATAGATGGCTTGTGTTATTAAGATAATCAACATTATCAGAAG 171
Db 61 TGAAGCAAGTGCAGATAGATGGCTTGTGTTATTAAGATAATCAACATTATCAGAAG 120

QY 172 AAGGACAAGGACTGAAGTGTTCAGAGAGTGCTTTTGGGTCTGGTTGTAAGAAGATCGGC 231
Db 121 AAGGACAAGGACTGAAGTGTTCAGAGAGTGCTTTTGGGTCTGGTTGTAAGAAGATCGGC 180

QY 232 TTGAAATTTACCACTGCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 291
Db 181 TTGAAATTTACCACTGCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 240

QY 292 AAGTCCAATATCAGATGCAATGATCGGAGCCTTCGCCATGTAAACATTGATCATCTTC 351
Db 241 AAGTCCAATATCAGATGCAATGATCGGAGCCTTCGCCATGTAAACATTGATCATCTTC 300

QY 352 ACGTGGCTGGTATCAGTCCACATATGCTCATTCGTTACCGGGCACTCCTGGACT 411
Db 301 ACGTGGCTGGTATCAGTCCACATATGCTCATTCGTTACCGGGCACTCCTGGACT 360

QY 412 CTCAGTTTAGTACCAGATGCAATGAGAACTGCTGCTTCATTTATGATCCCATAA 471
Db 361 CTCAGTTTAGTACCAGATGCAATGAGAACTGCTGCTTCATTTATGATCCCATAA 420

QY 472 AAATGCCCAAGGATCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAG 531
Db 421 AAATGCCCAAGGATCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAG 480

QY 532 TTTGTAAGAAAGGATTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGT 591
Db 481 TTTGTAAGAAAGGATTTTCCCTGAAGCATTTTCCCTGAA-CATTGAAAAAGCAAAATATCACCTTTGAGT 539

QY 592 ACATGTTTGAAGAAGTG-CCGATTTGTAATTAATAATTCACATCT--GATCAATGCTCTAA 648
Db 540 ACATGTTTGAAGAAGTGCCCGATTTGTAATTAATAATTCACATCTTTGATCAATGCTCTAA 599

QY 649 TGTGGAACTTGAAGAAGA 666
Db 600 TGTGGAACTTGAAGAAGA 617
```

```
RESULT 11
US-10-015-219-1299
; Sequence 1299, Application US/10015219
; Publication No. US20030008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015,219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1299
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 626, 640, 649, 655, 661, 670, 702, 712, 714, 720, 731, 744,
; LOCATION: 758, 764, 765
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-1299

Query Match      43.9%; Score 562.4; DB 14; Length 783;
Best Local Similarity 97.6%; Pred. No. 9.7e-121;
Matches 603; Conservative 0; Mismatches 11; Indels 4; Gaps 3;

QY 52 GCTCCACCGCGCGCGAGCAAGGCAAGCGCGCTCGGAGATTCAGCG 111
Db 1 GCTCCACCGCGCGCGAGCAAGGCAAGCGCGCTCGGAGATTCAGCG 60

QY 112 TGAAGCAAGTGCAGATAGATGGCTTGTGTTATTAAGATAATCAACATTATCAGAAG 171
Db 61 TGAAGCAAGTGCAGATAGATGGCTTGTGTTATTAAGATAATCAACATTATCAGAAG 120

QY 172 AAGGACAAGGACTGAAGTGTTCAGAGAGTGCTTTTGGGTCTGGTTGTAAGAAGATCGGC 231
Db 121 AAGGACAAGGACTGAAGTGTTCAGAGAGTGCTTTTGGGTCTGGTTGTAAGAAGATCGGC 180

QY 232 TTGAAATTTACCACTGCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 291
Db 181 TTGAAATTTACCACTGCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 240

QY 292 AAGTCCAATATCAGATGCAATGATCGGAGCCTTCGCCATGTAAACATTGATCATCTTC 351
Db 241 AAGTCCAATATCAGATGCAATGATCGGAGCCTTCGCCATGTAAACATTGATCATCTTC 300

QY 352 ACGTGGCTGGTATCAGTCCACATATGCTCATTCGTTACCGGGCACTCCTGGACT 411
Db 301 ACGTGGCTGGTATCAGTCCACATATGCTCATTCGTTACCGGGCACTCCTGGACT 360

QY 412 CTCAGTTTAGTACCAGATGCAATGAGAACTGCTGCTTCATTTATGATCCCATAA 471
Db 361 CTCAGTTTAGTACCAGATGCAATGAGAACTGCTGCTTCATTTATGATCCCATAA 420

QY 472 AAATGCCCAAGGATCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAG 531
Db 421 AAATGCCCAAGGATCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAG 480

QY 532 TTTGTAAGAAAGGATTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGT 591
Db 481 TTTGTAAGAAAGGATTTTCCCTGAA-CATTGAAAAAGCAAAATATCACCTTTGAGT 539

QY 592 ACATGTTTGAAGAAGTG-CCGATTTGTAATTAATAATTCACATCT--GATCAATGCTCTAA 648
Db 540 ACATGTTTGAAGAAGTGCCCGATTTGTAATTAATAATTCACATCTTTGATCAATGCTCTAA 599

QY 649 TGTGGAACTTGAAGAAGA 666
Db 600 TGTGGAACTTGAAGAAGA 617
```



```

RESULT 12
US-09-920-300A-168/c
; Sequence 168, Application US/09920300A
; Publication No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 13, 515, 516, 517, 565
; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-168

Query Match      43.9%; Score 562; DB 9; Length 567;
Best Local Similarity 99.1%; Pred. No. 1e-120;
Matches 562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 106 CAGCGTGAAGCAAGTCAGATAGATGGCTTGTGTTAAAGATAATCAAACTATTC 165
Db 567 CAGCGTGAAGCAAGTCAGATAGATGGCTTGTGTTAAAGATAATCAAACTATTC 508

Qy 166 AAGAAGAGGACAGGAAGTGAAGTTGTTCAAGAGTCTTTGGGTCTGGTTAGAG 225
Db 507 AAGAAGAGGACAGGAAGTGAAGTTGTTCAAGAGTCTTTGGGTCTGGTTAGAG 448

Qy 226 ATCGGCTTGAATATACCAACTGCTTTCTTCCCTCAGCACACAGAGGATGCTGACT 285
Db 447 ATCGGCTTGAATATACCAACTGCTTTCTTCCCTCAGCACACAGAGGATGCTGACT 388

Qy 286 TTGATGAAGTCCAATATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTTGATC 345
Db 387 TTGATGAAGTCCAATATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTTGATC 328

Qy 346 ATCTTCAGTGGGCTGGTATCAGTCCACATATGCTGCTCAATTTGTTACCGGGCACTCC 405
Db 327 ATCTTCAGTGGGCTGGTATCAGTCCACATATGCTGCTCAATTTGTTACCGGGCACTCC 268

Qy 406 TGGACTCTCAGTTTAGTTTACCAGCATGCCATTGAAGATCTGCTTCTCATTTATGATC 465
Db 267 TGGACTCTCAGTTTAGTTTACCAGCATGCCATTGAAGATCTGCTTCTCATTTATGATC 208

Qy 466 CCATAAAACTGCCCAAGGATCTCTCCTCACTAAAGGCATACAGACTGACTCCTAAACTGA 525
Db 207 CCATAAAACTGCCCAAGGATCTCTCCTCACTAAAGGCATACAGACTGACTCCTAAACTGA 148

Qy 526 TGGAGTTTGTAAAGAAAGGATTTTCCCTCCTGAAGCATTTGAAAGCAAAATATCACCT 585
Db 147 TGGAGTTTGTAAAGAAAGGATTTTCCCTCCTGAAGCATTTGAAAGCAAAATATCACCT 88

Qy 586 TTGAGTACATCTTGAAGAGAGTGGCGATTGTAATTAATAATTCACATCTGATCAATGTC 645
Db 87 TTGAGTACATCTTGAAGAGAGTGGCGATTGTAATTAATAATTCACATCTGATCAATGTC 28

Qy 646 TAATGTGGGAACCTTGAAGAGAGTGCAG 672
Db 27 TAATGTGGGAACCTTGAAGAGAGTGCAG 1

```

RESULT 13

```

US-10-033-528-168/c
; Sequence 168, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 13, 515, 516, 517, 565
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-168

Query Match      43.9%; Score 562; DB 13; Length 567;
Best Local Similarity 99.1%; Pred. No. 1e-120;
Matches 562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 106 CAGCGTGAAGCAAGTCAGATAGATGGCTTGTGTTAAAGATAATCAAACTATTC 165
Db 567 CAGCGTGAAGCAAGTCAGATAGATGGCTTGTGTTAAAGATAATCAAACTATTC 508

Qy 166 AAGAAGAGGACAGGAAGTGAAGTTGTTCAAGAGTCTTTGGGTCTGGTTAGAG 225
Db 507 AAGAAGAGGACAGGAAGTGAAGTTGTTCAAGAGTCTTTGGGTCTGGTTAGAG 448

Qy 226 ATCGGCTTGAATATACCAACTGCTTTCTTCCCTCAGCACACAGAGGATGCTGACT 285
Db 447 ATCGGCTTGAATATACCAACTGCTTTCTTCCCTCAGCACACAGAGGATGCTGACT 388

Qy 286 TTGATGAAGTCCAATATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTTGATC 345
Db 387 TTGATGAAGTCCAATATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTTGATC 328

Qy 346 ATCTTCAGTGGGCTGGTATCAGTCCACATATGCTGCTCAATTTGTTACCGGGCACTCC 405
Db 327 ATCTTCAGTGGGCTGGTATCAGTCCACATATGCTGCTCAATTTGTTACCGGGCACTCC 268

Qy 406 TGGACTCTCAGTTTAGTTTACCAGCATGCCATTGAAGATCTGCTTCTCATTTATGATC 465
Db 267 TGGACTCTCAGTTTAGTTTACCAGCATGCCATTGAAGATCTGCTTCTCATTTATGATC 208

Qy 466 CCATAAAACTGCCCAAGGATCTCTCCTCACTAAAGGCATACAGACTGACTCCTAAACTGA 525
Db 207 CCATAAAACTGCCCAAGGATCTCTCCTCACTAAAGGCATACAGACTGACTCCTAAACTGA 148

Qy 526 TGGAGTTTGTAAAGAAAGGATTTTCCCTCCTGAAGCATTTGAAAGCAAAATATCACCT 585
Db 147 TGGAGTTTGTAAAGAAAGGATTTTCCCTCCTGAAGCATTTGAAAGCAAAATATCACCT 88

Qy 586 TTGAGTACATCTTGAAGAGAGTGGCGATTGTAATTAATAATTCACATCTGATCAATGTC 645
Db 87 TTGAGTACATCTTGAAGAGAGTGGCGATTGTAATTAATAATTCACATCTGATCAATGTC 28

Qy 646 TAATGTGGGAACCTTGAAGAGAGTGCAG 672
Db 27 TAATGTGGGAACCTTGAAGAGAGTGCAG 1

```

RESULT 14  
US-10-099-926-168/c  
; Sequence 168, Application US/10099926

```
; Publication No. US20030166064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 13, 515, 516, 517, 565
; OTHER INFORMATION: n = A,T,C or G
US-10-099-926-168
```

```
Query Match 43.9%; Score 562; DB 15; Length 567;
Best Local Similarity 99.1%; Pred. No. 1e-120;
Matches 562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 106 CAGCCGTGAAGCAAGTGCAGATAGATGCGCTTGGTATTAAAGATAATCAACATATTC 165
Db |
QY 567 CAGCCGTGAAGCAAGTGCAGATAGATGCGCTTGGTATTAAAGATAATCAACATATTC 508
Db |
QY 166 AAGAAAGAGCAAGCAAGTGCAGATAGATGCGCTTGGTATTAAAGATAATCAACATATTC 225
Db |
QY 507 AAGAAAGAGCAAGCAAGTGCAGATAGATGCGCTTGGTATTAAAGATAATCAACATATTC 448
QY 226 ATCCGCTTGAATATACCAATGCTTTCTTCCCTCAGACACAGAGATGATGCTGACT 285
Db |
QY 447 ATCCGCTTGAATATACCAATGCTTTCTTCCCTCAGACACAGAGATGATGCTGACT 388
QY 286 TTGATGAAGTCCAAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAACATTTGATC 345
Db |
QY 387 TTGATGAAGTCCAAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAACATTTGATC 328
QY 346 ATCTTCAAGTGGCTGGTATCAGTCCATATGCTTCCCTCAGACACAGAGATGATGCTGACT 405
Db |
QY 327 ATCTTCAAGTGGCTGGTATCAGTCCATATGCTTCCCTCAGACACAGAGATGATGCTGACT 268
QY 406 TGGACTCTCAGTTTGTAGTTTACCAGATGCCATTTGAAGATCTGCTGCTCATTTATGATC 465
Db |
QY 267 TGGACTCTCAGTTTGTAGTTTACCAGATGCCATTTGAAGATCTGCTGCTCATTTATGATC 208
QY 466 CCATAAAAGTCCCAAGGATCTCTCTCACTAAAGCATACAGACTGACTCTCTAAACTGA 525
Db |
QY 207 CCATAAAAGTCCCAAGGATCTCTCTCACTAAAGCATACAGACTGACTCTCTAAACTGA 148
QY 526 TGGAGTTTGTAAAGAAAGATTTTCCCTCAGACATTTGAAGAAAGCAATATCACT 585
Db |
QY 147 TGGAGTTTGTAAAGAAAGATTTTCCCTCAGACATTTGAAGAAAGCAATATCACT 88
QY 586 TTGAGTACATGTTTGAAGAAAGTCCGATTTGAATTTAAATTTCAATCTGATCAATGTC 645
Db |
QY 87 TTGAGTACATGTTTGAAGAAAGTCCGATTTGAATTTAAATTTCAATCTGATCAATGTC 28
QY 646 TAATGTGGGAACCTTGAAGAAAGTCTAG 672
Db |
QY 27 TAATGTGGGAACCTTGAAGAAAGTCTAG 1
```

```
RESULT 15
US-09-969-034-1346/c
; Sequence 1346, Application US/09969034
; Publication No. US20040110668A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1346
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 552-563
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-1346
```

```
Query Match 40.2%; Score 514.4; DB 11; Length 571;
Best Local Similarity 98.4%; Pred. No. 1.3e-109;
Matches 541; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 45 TCTTCAGCTCCACCGCGG--CGCAGCAGGAAAGCAAGGCAAGGC--GGCTCGGGA 101
Db |
QY 550 TTTTCAGCTCCACCGCGGCGCAAGCAGGAAAGCAAGGCAAGGC--GGCTCGGGA 491
Db |
QY 102 GATTTCAGCGGTGAAGCAAGTGCAGATAGATGCGCTTGGTATTAAAGATAATCAACAT 161
Db |
QY 490 GATTTCAGCGGTGAAGCAAGTGCAGATAGATGCGCTTGGTATTAAAGATAATCAACAT 431
QY 162 TATCAAGAAAGAGCAAGGAACTGAAGTTCTTCAAGAGTGCCTTTGGGTCTGGTTGA 221
Db |
QY 430 TATCAAGAAAGAGCAAGGAACTGAAGTTCTTCAAGAGTGCCTTTGGGTCTGGTTGA 371
QY 222 GAAGATCGGCTTGAATTTACCACTGCTTCCCTTCCCTCAGACACAGAGATGATGCT 281
Db |
QY 370 GAAGATCGGCTTGAATTTACCACTGCTTCCCTTCCCTCAGACACAGAGATGATGCT 311
QY 282 GACTTTGATGAAGTCCCAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAACAT 341
Db |
QY 310 GACTTTGATGAAGTCCCAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAACAT 251
QY 342 GATCATCTTCAAGTGGCTGGTATCAGTCCACATATCTATGGCTCATTTGTTACCGGGA 401
Db |
QY 250 GATCATCTTCAAGTGGCTGGTATCAGTCCACATATCTATGGCTCATTTGTTACCGGGA 191
QY 402 CTCCTGACTCTCAGTTTGTAGTTTACCAGATGCCATTTGAAGAACTGCTGCTCATTTAT 461
Db |
QY 190 CTCCTGACTCTCAGTTTGTAGTTTACCAGATGCCATTTGAAGAACTGCTGCTCATTTAT 131
QY 462 GATCCCAATAAAAGTCCCAAGGATCTCTCTCACTAAAGCATACAGACTGACTCTCTAAA 521
Db |
QY 130 GATCCCAATAAAAGTCCCAAGGATCTCTCTCACTAAAGCATACAGACTGACTCTCTAAA 71
QY 522 CTGATGGAAGTCTTGAAGAAAGATTTTCCCTTGAAGCATTTGAAAAAGCAATATC 581
Db |
QY 70 CTGATGGAAGTCTTGAAGAAAGATTTTCCCTTGAAGCATTTGAAAAAGCAATATC 11
QY 582 ACCTTTGAGT 591
Db |
QY 10 ACCTTTGAGT 1
```



Search completed: January 5, 2005, 16:23:11  
Job time : 752 secs

---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 07:09:50 ; Search time 679 Seconds  
(without alignments)  
9895.816 Million cell updates/sec

Title: US-10-017-327-1  
Perfect score: 1280  
Sequence: 1 gaaagatggctccgcaag.....aaaaaaaaaaaaaaaaaaaa 1280

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1280	100.0	1280	3	AAA61875 Human eIF
2	1280	100.0	1280	6	ABA97211 Human mel
3	1280	100.0	1280	9	ACC85029 Human ant
4	1280	100.0	1280	10	ADG32712 Human DNA
5	1280	100.0	1280	10	ADG32712 Human DNA
6	1214.8	94.9	1337	3	Adf15978 Leukaemia
7	1139.4	89.0	2285	5	Adf15978 Human pro
8	1107	86.5	2455	10	Adc30341 Human nov
9	1045	81.6	1045	10	Adc30341 Human nov
10	908	70.9	908	10	Adc30341 Human nov
11	723.6	56.5	774	6	Abq55406 Human ova
12	619.4	48.4	621	4	Aas25348 Human ova
13	563.4	44.0	565	6	Abq55406 Human ova
14	562.4	43.9	783	4	Aas25348 Human ova
15	562.4	43.9	793	5	Aas25348 Human ova
16	562	43.9	567	6	Abk44617 cDNA enco
17	522	40.8	522	10	Adf79557 Leukaemia
18	514.4	40.2	571	6	Abq55406 Human ova
19	512.6	40.0	573	6	Abq55406 Human ova
20	501	39.1	723	4	Aas32948 DNA encod
21	476.2	37.2	760	10	Adc92366 Human col

22	471.4	36.8	518	3	AAC08361 Human sec
c 23	471.4	36.8	542	5	ADL41435 Human ova
c 24	451.6	35.3	651	6	ABQ60378 Human col
25	440.2	34.4	574	6	ABQ57787 Human col
26	438.8	34.3	597	6	ABQ57782 Human col
27	435	34.0	471	6	ABK16246 Human lun
28	435	34.0	471	10	ADB95509 Human lun
29	427.4	33.4	449	2	AAZ24520 Human lun
30	427.4	33.4	449	3	AAC65759 Human lun
31	427.4	33.4	449	6	ABL48978 Human lun
32	427.4	33.4	449	6	ABQ92164 Human lun
33	427.4	33.4	449	9	ADA28579 Human lun
34	427.4	33.4	449	10	ADE53539 Human lun
35	427.4	33.4	449	10	ADH36674 Human lun
36	427.4	33.4	449	12	ADM56477 Human lun
37	427.4	33.4	449	12	ADN89521 Human lun
c 38	426.2	33.3	557	6	ABQ57766 Human col
c 39	412.8	32.2	454	6	ABN96026 Gene #252
c 40	408.4	31.9	618	10	ADB51862 Primary r
c 41	401	31.3	564	6	ABQ58947 Human col
42	387.4	30.3	466	9	ACH42347 Human foe
43	362	28.3	1130	6	ABQ54653 Human ova
44	361.4	28.2	1149	2	AAX98024 Human sec
45	361.4	28.2	1149	9	ADA11588 Human CDN

ALIGNMENTS

RESULT 1  
AAA61875  
ID AAA61875 standard; CDNA; 1280 BP.  
AC AAA61875;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE Human eIF3-p40 coding sequence.  
XX  
KW eIF3-p40; eukaryotic translation initiation factor 3; p40 subunit; human;  
KW chromosome 8q23; breast carcinoma; prostate carcinoma; overexpression;  
KW gene amplification; oncogenesis; cancer; tumour; EST 346021;  
KW expressed sequence tag; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200036144-A2.  
XX  
PD 22-JUN-2000.  
XX  
PF 15-DEC-1999; 99WO-FI001039.  
XX  
PR 16-DEC-1998; 98FI-00002722.  
XX  
(FIIM-) FINNISH IMMUNOTECHNOLOGY LTD.  
XX  
PI Visakorpi T, Isola J, Nupponen N, Ovod V;  
XX  
DR WPI; 2000-431613/37.  
XX  
PT Detecting and modulating the expression of p40 subunits of eukaryotic  
PT translation initiation factor 3 for the diagnosis, prognosis and  
PT treatment of breast and prostate carcinomas.  
XX  
PS Example 6; Page 23-24; 24pp; English.  
XX  
PS The invention relates to novel methods and reagents for detecting and  
XX quantifying the expression of the p40 subunit of human eIF3-p40  
XX (eukaryotic translation initiation factor 3) and for treating conditions  
XX associated with inappropriate expression of eIF3-p40. eIF-p40 is  
XX associated with the oncogenesis of breast and prostate cancers. The eIF3-  
XX p40 gene is located on the long arm of chromosome 8 (8q23). The long arm  
XX of chromosome 8 is one of the most common regions that is amplified in

CC cancers of several organs, especially carcinomas of the breast and  
 CC prostate. eIF3-p40 is also overexpressed in a large fraction of breast  
 CC and prostate cancer, indicating that it is a target gene for 8q  
 CC amplification. eIF3-p40 protein, its functional variants or fragments,  
 CC and anti-eIF3-p40 antibodies may be used as diagnostic and therapeutic  
 CC agents for aggressive forms of cancer involving the inappropriate  
 CC expression of eIF3-p40. The present sequence represents the coding  
 CC sequence of the human eIF3-p40 gene, which is derived from EST (expressed  
 CC sequence tag) 346021  
 XX

SQ Sequence 1280 BP; 446 A; 273 C; 274 G; 287 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1280; DB 3; Length 1280;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-226;  
 Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAAAGATGGCGTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
 DB 1 GAAAGATGGCGTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
 QY 61 CCGCGCAGCAGGGAAGGCAAGGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
 DB 61 CCGCGCAGCAGGGAAGGCAAGGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
 QY 121 TGCAGATAGATGGCGTCTGTGTATTAAGATAATCAAAACATTATCAAGAAAGAGCAAG 180  
 DB 121 TGCAGATAGATGGCGTCTGTGTATTAAGATAATCAAAACATTATCAAGAAAGAGCAAG 180  
 QY 181 GAATGAAGTTGTTCAAGAGATGCTTTTGGGTCGTGGTTGAGAAGATCGGCTTGAATTA 240  
 DB 181 GAATGAAGTTGTTCAAGAGATGCTTTTGGGTCGTGGTTGAGAAGATCGGCTTGAATTA 240  
 QY 241 CCAACTGCTTTCCCTCCCTCAGCACACAGAGATGATGCTGATCTTCAAGTCCAAAT 300  
 DB 241 CCAACTGCTTTCCCTCCCTCAGCACACAGAGATGATGCTGATCTTCAAGTCCAAAT 300  
 QY 301 ATCAGATGGAATGATCGGAGCGCTTCGCCATGTAAACATTGATCATCTTCAAGTGGGCT 360  
 DB 301 ATCAGATGGAATGATCGGAGCGCTTCGCCATGTAAACATTGATCATCTTCAAGTGGGCT 360  
 QY 361 GGTATCAGTCCACATATCTATGGCTCACTGTTTACCGGGCACTCTCGGACTCTCAGTTTA 420  
 DB 361 GGTATCAGTCCACATATCTATGGCTCACTGTTTACCGGGCACTCTCGGACTCTCAGTTTA 420  
 QY 421 GTTACACAGATGCAATGGAAGATCTGCTGCTCTCAATTTATGATCCCAATAAACTGCC 480  
 DB 421 GTTACACAGATGCAATGGAAGATCTGCTGCTCTCAATTTATGATCCCAATAAACTGCC 480  
 QY 481 AAGATCTCTCTCACTAAAGGCATACAGACTGATCTCTAACTGATGGAAGTTTGTAAAG 540  
 DB 481 AAGATCTCTCTCACTAAAGGCATACAGACTGATCTCTAACTGATGGAAGTTTGTAAAG 540  
 QY 541 AAAAGGATTTTCCCTGAGCAATTGAAAAGCAAAATATCACTTTGAGTACATGTTTG 600  
 DB 541 AAAAGGATTTTCCCTGAGCAATTGAAAAGCAAAATATCACTTTGAGTACATGTTTG 600  
 QY 601 AAGAAGTGGCGATTGTAATTAATAATTCATCTGATCAATGCTCTAACTGTTGGAACTTG 660  
 DB 601 AAGAAGTGGCGATTGTAATTAATAATTCATCTGATCAATGCTCTAACTGTTGGAACTTG 660  
 QY 661 AAAAAGTCAAGTGTGAGATAAACAATGAATTTGCTCAGCCCTTGCAGCAATCAAT 720  
 DB 661 AAAAAGTCAAGTGTGAGATAAACAATGAATTTGCTCAGCCCTTGCAGCAATCAAT 720  
 QY 721 TGGGGAGAGATCTACAGATTGCTGATGGAAGATGATGATGATGATGATGATGATGATGAT 780  
 DB 721 TGGGGAGAGATCTACAGATTGCTGATGGAAGATGATGATGATGATGATGATGATGATGAT 780  
 QY 781 AATACACACATACATACAGAGATCTAGTAAACAACAGCAGCAGCAAAATCAGTATCAGC 840  
 DB 781 AATACACACATACATACAGAGATCTAGTAAACAACAGCAGCAGCAAAATCAGTATCAGC 840  
 QY 841 AGCGTCGCCAGGAGGAATATGACGCGCCAGAGCGAGAGAAACCCCGCTCCCTGAGG 900

DB 841 AGCGTCGCCAGGAGGAATATGACGCGCCAGAGCGGAGAAACCCCGCTCCCTGAGG 900  
 QY 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCGGCTGCGAGATGGACTCGCTGCTCA 960  
 DB 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCGGCTGCGAGATGGACTCGCTGCTCA 960  
 QY 961 TTGCAGGCGAGATAAACAATCTTACTGCGCAACAATCAAGAGTTCACTGCGCAAACTTAG 1020  
 DB 961 TTGCAGGCGAGATAAACAATCTTACTGCGCAACAATCAAGAGTTCACTGCGCAAACTTAG 1020  
 QY 1021 GCAAGCTCTTTCATGCGCCAGGCTCTTCAAGAATACAACTAAGAAAGGAGTTTCCA 1080  
 DB 1021 GCAAGCTCTTTCATGCGCCAGGCTCTTCAAGAATACAACTAAGAAAGGAGTTTCCA 1080  
 QY 1081 GAAAAGAGTTTAAACATGAACTCTTGAAGTCAACACAGGCGCACTCTTGAAGAAATATAT 1140  
 DB 1081 GAAAAGAGTTTAAACATGAACTCTTGAAGTCAACACAGGCGCACTCTTGAAGAAATATAT 1140  
 QY 1141 TTGCATATTTGAAAAGCAGAGAGATTTCTTTAGTGTCTATGCGGATTTTGGCTATAACAG 1200  
 DB 1141 TTGCATATTTGAAAAGCAGAGAGATTTCTTTAGTGTCTATGCGGATTTTGGCTATAACAG 1200  
 QY 1201 TGTCTTTCTAGCATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
 DB 1201 TGTCTTTCTAGCATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
 QY 1261 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1280  
 DB 1261 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1280  
 RESULT 2  
 ID ABA97211 standard; cDNA; 1280 BP.  
 XX ABA97211;  
 AC ABA97211;  
 XX 07-MAY-2002 (first entry)  
 DE Human melanoma antigen eukaryotic initiation factor 3 (eIF3) cDNA.  
 KW Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cyklostatic; immunomodulator; immune effector cell;  
 XX anti-cancer; vaccine; db.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 6..1064  
 FT /\*tag= a  
 FT /product= "eIF3"  
 XX WO200192307-A2.  
 PN 06-DEC-2001.  
 PD 30-MAY-2001; 2001WO-US017456.  
 XX 31-MAY-2000; 2000US-0209391P.  
 PR 17-AUG-2000; 2000US-0226258P.  
 PR 20-DEC-2000; 2000US-0257008P.  
 XX (GENZ ) GENZYME CORP.  
 PA Nicolette CA;  
 PI WPI; 2002-139606/18.  
 DR P-PSDB; ABB08360.  
 DR New therapeutic compounds useful against human ovarian cancer, for  
 PT modulating immune response in a subject, and for generating antibodies  
 PT that specifically recognize and bind to these molecules.

XX PS Disclosure; Page 63; 68pp; English.

CC The invention relates to novel therapeutic compounds, that are designed to enhance binding to MHC molecules and to enhance immunoregulatory properties relative to their natural counterparts. The activity of the compounds of the invention may be described as cytostatic and immunomodulatory. The compounds are useful against human ovarian cancer, for modulating immune response in a subject, and for generating antibodies that specifically recognize and bind to these molecules.

CC Compositions comprising the compounds are useful as components of anti-cancer vaccines and to expand immune effector cells that are specific for cells characterized by expression of antigen eIF3 (melanoma antigen eukaryotic initiation factor). The peptides or polypeptides conjugated to a detectable agent may be used in diagnostic procedures, such as in the detection and purification of antibodies, and as immunogens for production of antibodies. The polynucleotides can be used as primers for detecting genes or gene transcripts expressed in APC to confirm transduction of the polynucleotides into host cells. The current sequence CC represents the human melanoma antigen eukaryotic initiation factor 3 (eIF3) cDNA

XX SQ Sequence 1280 BP; 446 A; 273 C; 274 G; 287 T; 0 U; 0 Other;

Query Match 100.0%; Score 1280; DB 6; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 1.9e-226;  
Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGATGGCTCCGCGAAGGATACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Db 1 GAAAGATGGCTCCGCGAAGGATACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60

Qy 61 CCGCGCAGCAGGGAAGGCAAGCGGCTCGGGAGATTTCAGCGGTGAAGCAAG 120  
Db 61 CCGCGCAGCAGGGAAGGCAAGCGGCTCGGGAGATTTCAGCGGTGAAGCAAG 120

Qy 121 TGCAGATAGATGGCTGTGGTATTAAAGATAATCAAAATTATCAAGAAAGGACAAAG 180  
Db 121 TGCAGATAGATGGCTGTGGTATTAAAGATAATCAAAATTATCAAGAAAGGACAAAG 180

Qy 181 GAATGAAGTTGTTCAAGGAGTGCTTTGGGTCTGGTTGTGAAGATCGGCTTGAATTA 240  
Db 181 GAATGAAGTTGTTCAAGGAGTGCTTTGGGTCTGGTTGTGAAGATCGGCTTGAATTA 240

Qy 241 CCAATGCTTTCTTCCCTCAGCACACAGAGGATGATGCTGACTTTTGATGAAGTCCAAT 300  
Db 241 CCAATGCTTTCTTCCCTCAGCACACAGAGGATGATGCTGACTTTTGATGAAGTCCAAT 300

Qy 301 ATCAGATGGAATGATGCGAGCCTTCGCCATGTAACATTTGATCATCTTCACGTGGGCT 360  
Db 301 ATCAGATGGAATGATGCGAGCCTTCGCCATGTAACATTTGATCATCTTCACGTGGGCT 360

Qy 361 GGTATCAGTCACATACATGCTGCTCATTCGTTACCGGGCACTCTGAGCTCTCAGTTTA 420  
Db 361 GGTATCAGTCACATACATGCTGCTCATTCGTTACCGGGCACTCTGAGCTCTCAGTTTA 420

Qy 421 GTTACCAGCATGCCATTGAAGAATCTGCTTCTCATTTATGATGCCATAAAACTGCC 480  
Db 421 GTTACCAGCATGCCATTGAAGAATCTGCTTCTCATTTATGATGCCATAAAACTGCC 480

Qy 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAAGTTGTTAAG 540  
Db 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAAGTTGTTAAG 540

Qy 541 AAAAGGATTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTG 600  
Db 541 AAAAGGATTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTG 600

Qy 601 AAGAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGTCTTAATGTGGAACTTG 660  
Db 601 AAGAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGTCTTAATGTGGAACTTG 660

Qy 661 AAAAGAGTCAGCTGTTGCAGATAAAACATGAATTGCTCAGCGCTTGCCAGCAGCAATCAT 720

Db 661 AAAAGAGTCAGCTGTTGCAGATAAAACATGAATTTGCTCAGCGCTTGCCAGCAGCAATCAT 720

Qy 721 TGGGGAAGAACTCTACAGTTGCTGATGGACAGAGTGGATGAATGAGCCAAAGATATAGTTA 780  
Db 721 TGGGGAAGAACTCTACAGTTGCTGATGGACAGAGTGGATGAATGAGCCAAAGATATAGTTA 780

Qy 781 AATACAAACACATACATGAGGAATACTAGTAAACACAGCAGCAGAGCAAAACATCAGTATCAGC 840  
Db 781 AATACAAACACATACATGAGGAATACTAGTAAACACAGCAGCAGAGCAAAACATCAGTATCAGC 840

Qy 841 AGCTCGCCAGCAGAGAAATATGAGCCGCGCAGAGCCGAGGAGAAACCCCGCTCCCTGAGG 900  
Db 841 AGCTCGCCAGCAGAGAAATATGAGCCGCGCAGAGCCGAGGAGAAACCCCGCTCCCTGAGG 900

Qy 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCGCGCTGCCAGGATGAGCTCGCTGCTCA 960  
Db 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCGCGCTGCCAGGATGAGCTCGCTGCTCA 960

Qy 961 TTGCAGGCCAGATAAACACTTTACTGCCAGAACATCAAGGAGTTCACTGCCCCCAAACTTAG 1020  
Db 961 TTGCAGGCCAGATAAACACTTTACTGCCAGAACATCAAGGAGTTCACTGCCCCCAAACTTAG 1020

Qy 1021 GCAAGCTCTTCATGSCCGCAGGCTCTTCAAGAAATCAACAACTAAAGAAAGGAAAGTTTCCA 1080  
Db 1021 GCAAGCTCTTCATGSCCGCAGGCTCTTCAAGAAATCAACAACTAAAGAAAGGAAAGTTTCCA 1080

Qy 1081 GAAAGAGAGTTAAACATCAACTCTTGAAGTCAACAGCGGCAACTCTTTGGAAAGAAATATAT 1140  
Db 1081 GAAAGAGAGTTAAACATCAACTCTTGAAGTCAACAGCGGCAACTCTTTGGAAAGAAATATAT 1140

Qy 1141 TTGCATATTGAAAAGCACAGAGGATTTCTTTAGTGTGTCATTTGCCGATTTTGGCTATAACAG 1200  
Db 1141 TTGCATATTGAAAAGCACAGAGGATTTCTTTAGTGTGTCATTTGCCGATTTTGGCTATAACAG 1200

Qy 1201 TGTCTTTCTAGCCATAATAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260  
Db 1201 TGTCTTTCTAGCCATAATAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260

Qy 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
Db 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

RESULT 3  
ACC85029  
ID ACC85029 standard; cDNA; 1280 BP.  
XX AC ACC85029;  
XX AC  
XX AC  
DT 13-OCT-2003 (first entry)  
XX  
DE Human antigen eukaryotic translation initiation factor 3 (eIF3) cDNA.  
XX Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
KW cytosolic; gene therapy; human; gene; ss.  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 6..1064  
FT /\*tag= a  
FT /product= "eIF3"  
FT /note= "eukaryotic translation initiation factor 3"  
XX WO2003050543-A1.  
FN  
XX  
PD 19-JUN-2003.  
XX  
PF 05-DEC-2001; 2001WO-US047997.  
XX  
PR 05-DEC-2001; 2001WO-US047997.  
XX

(GENZ ) GENZYME CORP.

Nicolette CA;

WPI; 2003-532936/50.

P-PSDB; ABR82212.

Aiding in the diagnosis of a neoplastic condition, useful for treating cancer and related malignancies comprises determining the amount of expression of an eIF3 protein in a test sample isolated from the cell or tissue.

Claim 5; Page 71; 77pp; English.

The invention relates to aiding in the diagnosis of a neoplastic condition or susceptibility to a neoplastic condition of an animal cell or tissue. The method involves determining the amount of expression of an eukaryotic translation initiation factor 3 (eIF3) protein in a test sample isolated from the cell or tissue, and diagnosing a neoplastic condition or susceptibility to a neoplastic condition based on the amount of expression of the eIF3 protein. The methods, compounds and kits are useful in therapeutics, diagnostic and screening methods for human cancer and related malignancies, e.g. ovarian, breast, lung, colon, prostate, pancreatic or gastrointestinal cancer, or melanoma. The present sequence represents the human antigen eIF3 encoding cDNA

Sequence 1280 BP; 446 A; 273 C; 274 G; 287 T; 0 U; 0 Other;

Query Match 100.0%; Score 1280; DB 9; Length 1280;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-226;  
 Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GAAAGATGGCTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
 1 GAAAGATGGCTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
 61 CCGCGCAGCAGGGAAGGCAAGGAAAGCGGCTCGGAGATTCAGCGTGAAGCAAG 120  
 61 CCGCGCAGCAGGGAAGGCAAGGAAAGCGGCTCGGAGATTCAGCGTGAAGCAAG 120  
 121 TGCAGATAGATGGCTTTGTGTATTAAGATATCAACATTATCAAGAAAGGACCAAG 180  
 121 TGCAGATAGATGGCTTTGTGTATTAAGATATCAACATTATCAAGAAAGGACCAAG 180  
 181 GAATGAAGTGTTCGAAGGATGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAATTA 240  
 181 GAATGAAGTGTTCGAAGGATGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAATTA 240  
 241 CCAACTGCTTTCCCTTCCCTCAGCACACAGAGGATGCTGACTTTGATGAAGTCCAAT 300  
 241 CCAACTGCTTTCCCTTCCCTCAGCACACAGAGGATGCTGACTTTGATGAAGTCCAAT 300  
 301 ATCAGATGGAATGATCGGAGGCTTCGCCATGTAAACATTGATCATCTTCAGTGGCT 360  
 301 ATCAGATGGAATGATCGGAGGCTTCGCCATGTAAACATTGATCATCTTCAGTGGCT 360  
 361 GGTATCAGTCCACATATAGGCTCATTCGTACCGGGCACTTCCTGGACTCTCAGTTTA 420  
 361 GGTATCAGTCCACATATAGGCTCATTCGTACCGGGCACTTCCTGGACTCTCAGTTTA 420  
 421 GTTACCAGATGCAATGAGATCTGCTTCTGATTTATGATCCATAAAACATGCGCC 480  
 421 GTTACCAGATGCAATGAGATCTGCTTCTGATTTATGATCCATAAAACATGCGCC 480  
 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTAAACTGATGGAAGTTGTAAG 540  
 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTAAACTGATGGAAGTTGTAAG 540  
 541 AAAGAGATTTTCCCTGAAGCATTCGAAAGCAAAATATCACTTTGAGTACATGTTTG 600  
 541 AAAGAGATTTTCCCTGAAGCATTCGAAAGCAAAATATCACTTTGAGTACATGTTTG 600  
 601 AAGAAGTCCGATGTGAATTAATAATTCACATCTGATCAATGTCTTGAATGGGAACCTG 660

Db 601 AAGAAGTCCGATGTGAATTAATAATTCACATCTGATCAATGTCTTAATGTGGAACTTG 660  
 Qy 661 AAAAGAAGTCAGCTGTTGCAGATAAAACATGAATTTGCTCAGCCTTGCAGCAGCAATCAAT 720  
 Db 661 AAAAGAAGTCAGCTGTTGCAGATAAAACATGAATTTGCTCAGCCTTGCAGCAGCAATCAAT 720  
 Qy 721 TGGGGAAGATCTACAGTTGCTGTGATGAGCAGAGTGGATGAAATGAGCCAGATATAGTTA 780  
 Db 721 TGGGGAAGATCTACAGTTGCTGTGATGAGCAGAGTGGATGAAATGAGCCAGATATAGTTA 780  
 Qy 781 AATACAACACATACATGAGGAATACTAGTAAACAACAGCAGCAGCAACATCAGTATCAGC 840  
 Db 781 AATACAACACATACATGAGGAATACTAGTAAACAACAGCAGCAGCAACATCAGTATCAGC 840  
 Qy 841 AGCGTCGCCAGCAGGAGAATATGCGCGCCAGCGCAGCGAGAGAACCCCGCTCCCTGAGG 900  
 Db 841 AGCGTCGCCAGCAGGAGAATATGCGCGCCAGCGCAGCGAGAGAACCCCGCTCCCTGAGG 900  
 Qy 901 AGGACCTGTCCAAACTCTTCAACCCACACAGCGCGCTGCCAGGATGGAATCGCTGCTCA 960  
 Db 901 AGGACCTGTCCAAACTCTTCAACCCACACAGCGCGCTGCCAGGATGGAATCGCTGCTCA 960  
 Qy 961 TTGCAGGCCAGATAAACAATCTTACTGCCAGAACATCAAGGATTCAGTCCCAAACTTAG 1020  
 Db 961 TTGCAGGCCAGATAAACAATCTTACTGCCAGAACATCAAGGATTCAGTCCCAAACTTAG 1020  
 Qy 1021 GCAAGCTTTCATGCGCCAGGCTCTTCAAGAATACAACAATTAAGAAAGGAGTTTCCA 1080  
 Db 1021 GCAAGCTTTCATGCGCCAGGCTCTTCAAGAATACAACAATTAAGAAAGGAGTTTCCA 1080  
 Qy 1081 GAAAAGAGTTAAACATGAATCTTGAAGTTCACACAGCGCACTCTTGAAGTTCACACAGGCAATATAT 1140  
 Db 1081 GAAAAGAGTTAAACATGAATCTTGAAGTTCACACAGCGCACTCTTGAAGTTCACACAGGCAATATAT 1140  
 Qy 1141 TTGCATATTGAAAAGCAGACAGAGATTTCTTTAGTGTCTATTGCGGATTTTGGCTATAACAG 1200  
 Db 1141 TTGCATATTGAAAAGCAGACAGAGATTTCTTTAGTGTCTATTGCGGATTTTGGCTATAACAG 1200  
 Qy 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
 Db 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
 Qy 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
 Db 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

RESULT 4

ADG32712

ID ADG32712 standard; DNA; 1280 BP.

XX ADG32712;

XX 26-FEB-2004 (first entry)

XX Human DNA differentially expressed in patients with SLE SeqID36.

XX human; ds; autoimmune; chronic inflammatory disease; SLE;

XX systemic lupus erythematosus; rheumatoid arthritis; cholecystitis;

XX Sjogren's disease; CREST syndrome; scleroderma; ankylosing spondylitis;

XX ulcerative colitis; primary sclerosing cholangitis; appendicitis;

XX Homo sapiens.

XX WO2003090694-A2.

XX 06-NOV-2003.

XX 24-APR-2003; 2003WO-US013015.

XX 24-APR-2002; 2002US-00131827.

XX (EXPR-) EXPRESSION DIAGNOSTICS INC.  
XX Wohlgenuth J, Fry K, Woodward R, Ly N;  
XX WPI; 2003-877243/81.  
XX  
XX Diagnosing or monitoring autoimmune and chronic inflammatory diseases,  
XX such as rheumatoid arthritis, systemic lupus erythematosus, ulcerative  
XX colitis, psoriasis and asthma by detecting the expression level of one or  
XX more genes.  
XX  
XX Claim 18; SEQ ID NO 36; 877pp; English.  
XX  
XX This invention relates to novel methods for diagnosing and monitoring  
XX autoimmune and chronic inflammatory diseases. Specifically, it refers to  
XX the identification of genes that have a clinical utility as diagnostic  
XX tools for the management of, in particular, patients with systemic lupus  
XX erythematosus (SLE) or rheumatoid arthritis (RA). Accordingly, the  
XX present invention describes a method for determining the levels of  
XX multiple differentially expressed genes of a patient, in a concerted  
XX manner, in order to achieve an improved diagnostic assay with sensitivity  
XX and specificity for the disease in question. As such, these genes are  
XX useful for the diagnosis of various other inflammatory disorders  
XX including cholecystitis, Sjogren's disease, CREST syndrome, scleroderma,  
XX ankylosing spondylitis, ulcerative colitis, primary sclerosing  
XX cholangitis, appendicitis, diverticulitis, and primary biliary sclerosis.  
XX This polynucleotide is a DNA sequence representing human mRNA that is  
XX differentially expressed in patients with SLE, used in an exemplification  
XX of the invention.

XX SQ Sequence 1280 BP; 446 A; 273 C; 274 G; 287 T; 0 U; 0 Other;

Query Match 100.0%; Score 1280; DB 10; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 1.9e-226;  
Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGATGGCGTCCGCAAGGAAGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Db 1 GAAAGATGGCGTCCGCAAGGAAGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Qy 61 CCGCGCAGCAGGGAAGCAAGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
Db 61 CCGCGCAGCAGGGAAGCAAGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
Qy 121 TGCAGATAGTGGCTTGTGGTATTAAGATAATCAACATTATCAAGAAAGGACAAG 180  
Db 121 TGCAGATAGTGGCTTGTGGTATTAAGATAATCAACATTATCAAGAAAGGACAAG 180  
Qy 181 GAACTGAAGTTGTTCAAGGAGTCTTTGGGCTCTGGTGTGTAAGAGATCGGCTTGAATTA 240  
Db 181 GAACTGAAGTTGTTCAAGGAGTCTTTGGGCTCTGGTGTGTAAGAGATCGGCTTGAATTA 240  
Qy 241 CCAACTGCTTTCTTCCCTCAGCACACAGAGGATGATGCTGATTTGATGAAGTCCAAT 300  
Db 241 CCAACTGCTTTCTTCCCTCAGCACACAGAGGATGATGCTGATTTGATGAAGTCCAAT 300  
Qy 301 ATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTGATCATCTTCACGTGGGCT 360  
Db 301 ATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTGATCATCTTCACGTGGGCT 360  
Qy 361 GGTATCAGTCACATACATGCTTCTGCTTACCGGGCACTCTGAGCTCTCAGTTTA 420  
Db 361 GGTATCAGTCACATACATGCTTCTGCTTACCGGGCACTCTGAGCTCTCAGTTTA 420  
Qy 421 GTTACAGCATGCCATTGAAGAATCTGCTTCTCAATTTATGATGCCATAAAACTGCC 480  
Db 421 GTTACAGCATGCCATTGAAGAATCTGCTTCTCAATTTATGATGCCATAAAACTGCC 480  
Qy 481 AAGGATCTCTCTCAATAAGGCATACAGACTGACTCTTAACTGATGGAGTTTGTAAAG 540  
Db 481 AAGGATCTCTCTCAATAAGGCATACAGACTGACTCTTAACTGATGGAGTTTGTAAAG 540

Qy 541 AAAAGGATTTTCCCTGAAAGCATTGAAAAAGCAAAATATACCTTTGAGTACATGTTTG 600  
Db 541 AAAAGGATTTTCCCTGAAAGCATTGAAAAAGCAAAATATACCTTTGAGTACATGTTTG 600  
Qy 601 AAGBAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGCTTAATGTGGAACTTG 660  
Db 601 AAGBAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGCTTAATGTGGAACTTG 660  
Qy 661 AAAAGAAAGTCAAGCTTTGCGAGATAAATGCTCAGCTTTGCCAGCAGCAATCATTT 720  
Db 661 AAAAGAAAGTCAAGCTTTGCGAGATAAATGCTCAGCTTTGCCAGCAGCAATCATTT 720  
Qy 721 TGGGGAAGAACTTACAGTTGCTGATGACAGATGGAATGAATGAGCCAGATATAGTTA 780  
Db 721 TGGGGAAGAACTTACAGTTGCTGATGACAGATGGAATGAATGAGCCAGATATAGTTA 780  
Qy 781 AATACACACATACATCAGGAATACTAGTAACCAACAGCAGCAGCAAACTCAGTATCAGC 840  
Db 781 AATACACACATACATCAGGAATACTAGTAACCAACAGCAGCAGCAAACTCAGTATCAGC 840  
Qy 841 AGCTCCGCGCAGGAGAAATATGCGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAGG 900  
Db 841 AGCTCCGCGCAGGAGAAATATGCGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAGG 900  
Qy 901 AGGACCTGTCCAACTCTTCAAAACCAACAGCAGCCGCTGCGAGATGAGTCTGCTGCTCA 960  
Db 901 AGGACCTGTCCAACTCTTCAAAACCAACAGCAGCCGCTGCGAGATGAGTCTGCTGCTCA 960  
Qy 961 TTGCAGGCCAGATAAACAACATCTTCTGCGAGACATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Db 961 TTGCAGGCCAGATAAACAACATCTTCTGCGAGACATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Qy 1021 GCAAGCTCTTCTATGCGCCAGGCTCTTCAAGAATAACAACAACTAAGAAAAAGGAGTTTCCA 1080  
Db 1021 GCAAGCTCTTCTATGCGCCAGGCTCTTCAAGAATAACAACAACTAAGAAAAAGGAGTTTCCA 1080  
Qy 1081 GAAAGAAAGTTAAATGAACTCTTTGAAAGTCAACAGGCGCAACTCTTGGAAAGAAATATAT 1140  
Db 1081 GAAAGAAAGTTAAATGAACTCTTTGAAAGTCAACAGGCGCAACTCTTGGAAAGAAATATAT 1140  
Qy 1141 TTGCATATTTGAAGACACAGAGGATTTCTTTAGTGTCTATTTGCCGATTTTGGCTATAACAG 1200  
Db 1141 TTGCATATTTGAAGACACAGAGGATTTCTTTAGTGTCTATTTGCCGATTTTGGCTATAACAG 1200  
Qy 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
Db 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
Qy 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
Db 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

RESULT 5  
ADP81506  
ID ADP81506 standard; DNA; 1280 BP.  
XX AC ADP81506;  
XX AC ADP81506;  
XX DT 26-FEB-2004 (first entry)  
XX DE Leukaemia-related DNA sequence #2062.  
XX KW Cytostatic; Gene therapy; leukaemia; ss.  
XX OS Unidentified.  
XX XX  
XX FN WO2003039443-A2.  
XX XX  
XX PD 15-MAY-2003.  
XX XX  
XX PF 04-NOV-2002; 2002WO-EP012303.  
XX XX

PR 05-NOV-2001; 2001EP-00126244.  
 PR 30-APR-2002; 2002EP-00009758.  
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 PA (UJLU-) UNIV LUDWIG MAXIMILIANS.  
 PA (HAPE/) HAPERLACH T.  
 PA (SCHO/) SCHOCH C.  
 PA (KERN/) KERN W.  
 XX  
 PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;  
 PI Eils R, Brors B, Mergenthaler S;  
 XX WPI; 2003-505037/47.  
 DR  
 XX  
 PT Determining the subtype of leukemia cells and whether a patient sample  
 PT contains leukemia cells or other cells, useful for treating leukemia,  
 PT comprises determining the expression profile of a group of markers in a  
 PT patient sample.  
 XX  
 PS Disclosure; SEQ ID NO 2062; 2938pp; English.  
 CC  
 CC The present invention relates to a method (M1) for determining the  
 CC subtype of leukaemia cells and whether a patient sample contains  
 CC leukaemia cells. The method comprises determining the expression profile  
 CC of a group of markers in a patient sample. The method is useful for  
 CC determining the presence of leukaemia cells, its types or subtypes, and  
 CC for the preparation of a medicament for treating leukaemia.  
 XX  
 SQ Sequence 1280 BP; 446 A; 273 C; 274 G; 287 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1280; DB 10; Length 1280;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-226;  
 Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GAAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGTCCACCG 60  
 1 GAAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGTCCACCG 60  
 61 CCGCGCAGCAGGAGGAAAGGCAAGCGGCTCGGAGATTTCAGCCGTCAAGCAAG 120  
 61 CCGCGCAGCAGGAGGAAAGGCAAGCGGCTCGGAGATTTCAGCCGTCAAGCAAG 120  
 121 TGCAGATAGATGGCTTGGTATTAAGATATCAAAATTATCAAGAGAGGACAAG 180  
 121 TGCAGATAGATGGCTTGGTATTAAGATATCAAAATTATCAAGAGAGGACAAG 180  
 181 GAACCTGAAGTTGTTCAAGGAGTCTTTGGGCTGGTTGAGAGATCGCTTGAATTA 240  
 181 GAACCTGAAGTTGTTCAAGGAGTCTTTGGGCTGGTTGAGAGATCGCTTGAATTA 240  
 241 CCAACTGCTTTTCCCTCAGCAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
 241 CCAACTGCTTTTCCCTCAGCAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
 301 ATCAGATGGAATGATGCGGAGCTTGGCCATGTAACATTTGATCATCTTACGTTGGGCT 360  
 301 ATCAGATGGAATGATGCGGAGCTTGGCCATGTAACATTTGATCATCTTACGTTGGGCT 360  
 361 GGATCAGTCCACATGCTGCTCATCTGTTACCGGCGACTCTCGGACTCTCAGTTTA 420  
 361 GGATCAGTCCACATGCTGCTCATCTGTTACCGGCGACTCTCGGACTCTCAGTTTA 420  
 421 GTTACAGCATGCTGATGAAGATCTGCTGCTTCTTATTTATGATCCATAAAAGTGGCC 480  
 421 GTTACAGCATGCTGATGAAGATCTGCTGCTTCTTATTTATGATCCATAAAAGTGGCC 480  
 481 AAGGATCTCTCCTCACTAAAGGCATACAGACTGACTGACTGACTGAGGAGTTTGAAG 540  
 481 AAGGATCTCTCCTCACTAAAGGCATACAGACTGACTGACTGACTGAGGAGTTTGAAG 540  
 541 AAAAGGATTTTCCCTGAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTG 600  
 541 AAAAGGATTTTCCCTGAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTG 600

RESULT 6  
 AAF15978  
 ID AAF15978 standard; cDNA; 1337 BP.  
 XX  
 AC AAF15978;  
 XX  
 DT 13-MAR-2001 (first entry)  
 XX  
 DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:413.  
 DE  
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
 KW vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 XX WO200055174-A1.  
 FN  
 XX  
 PD 21-SEP-2000.  
 XX

QY 601 AAGAAGTCCGATTTGTAATTAATAATTCATCTGATCAATGTCCTTAATGTGGAACTTG 660  
 DB |||||  
 601 AAGAAGTCCGATTTGTAATTAATAATTCATCTGATCAATGTCCTTAATGTGGAACTTG 660  
 QY 661 AAAAAGATCAGCTGTTGAGATTAACATGATGCTGAGCTTGCAGCAGCAATCAATT 720  
 DB |||||  
 661 AAAAAGATCAGCTGTTGAGATTAACATGATGCTGAGCTTGCAGCAGCAATCAATT 720  
 QY 721 TGGGGAAGATCTTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCCAAGATATAGTTA 780  
 DB |||||  
 721 TGGGGAAGATCTTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCCAAGATATAGTTA 780  
 QY 781 AATACAAACATACATGAGGAATCTAGTAAACACAGCAGCAGCAAAATCATAGTATCAGC 840  
 DB |||||  
 781 AATACAAACATACATGAGGAATCTAGTAAACACAGCAGCAGCAAAATCATAGTATCAGC 840  
 QY 841 AGCGTCGCCAGCAGGAGGAATATGCGCGCCAGCAGCGGAGGAAACCCCGCTCCCTGAGG 900  
 DB |||||  
 841 AGCGTCGCCAGCAGGAGGAATATGCGCGCCAGCAGCGGAGGAAACCCCGCTCCCTGAGG 900  
 QY 901 AGGACCTGTCCTCAAACTCTTCAAAACCAACACAGCCCGCTGCCAGGATGGAATCGCTGCTCA 960  
 DB |||||  
 901 AGGACCTGTCCTCAAACTCTTCAAAACCAACACAGCCCGCTGCCAGGATGGAATCGCTGCTCA 960  
 QY 961 TTGCGAGCCAGATTAACACTTACTGCCAGACATCAAGGAGTTCAGTCCCAAACTTAG 1020  
 DB |||||  
 961 TTGCGAGCCAGATTAACACTTACTGCCAGACATCAAGGAGTTCAGTCCCAAACTTAG 1020  
 QY 1021 GCAAGCTCTTCAATGCGCCAGGCTCTTCAAGATAACAACAACTAAGAAAGGAGTTTCCA 1080  
 DB |||||  
 1021 GCAAGCTCTTCAATGCGCCAGGCTCTTCAAGATAACAACAACTAAGAAAGGAGTTTCCA 1080  
 QY 1081 GAAAAGATTTAACTGATGAACTCTTGAAGTCAACACAGGCAACTCTTGGAGAAATATAT 1140  
 DB |||||  
 1081 GAAAAGATTTAACTGATGAACTCTTGAAGTCAACACAGGCAACTCTTGGAGAAATATAT 1140  
 QY 1141 TTGCAATTTGAAAAGCAGAGGATTTCTTTAGTGTCTATTGCCGATTTTGGCTATAACAG 1200  
 DB |||||  
 1141 TTGCAATTTGAAAAGCAGAGGATTTCTTTAGTGTCTATTGCCGATTTTGGCTATAACAG 1200  
 QY 1201 TGTCTTTCTAGCCATAATAATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260  
 DB |||||  
 1201 TGTCTTTCTAGCCATAATAATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260  
 QY 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
 DB |||||  
 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

Human prostate cancer antigen nucleotide sequence SEQ ID NO:413.  
 Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
 vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;  
 antibacterial; gene therapy; neural; immune; reproductive; renal;  
 gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 wound; infectious disease; ss.

Homo sapiens.  
 WO200055174-A1.  
 21-SEP-2000.





PF 21-MAR-2001; 2001WO-US009126.  
XX 21-MAR-2000; 2000US-0191031P.  
PR 25-MAY-2000; 2000US-0207124P.  
PR 15-JUN-2000; 2000US-0211940P.  
PR 07-JUL-2000; 2000US-0216820P.  
PR 25-JUL-2000; 2000US-0220661P.  
PR 21-DEC-2000; 2000US-0257672P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA Lee J, Lillie J;  
PI WPI; 2001-611502/70.  
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
PT cancer cells as compared to their normal non-cancerous ovarian cells are  
PT used to characterize stage, grade, histological type of ovarian cancer.  
XX Disclosure; SEQ ID NO 20501; 106pp; English.  
XX The invention relates to nucleic acid markers which are overexpressed in  
CC ovarian cancer cells as compared to their expression in normal (i.e. non-  
CC cancerous) ovarian cells. The invention also relates to polypeptides  
CC encoded by the markers, antibodies that selectively bind to the  
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
CC of developing ovarian cancer involving inhibiting expression of a gene  
CC corresponding to a marker of the invention and a method of treating a  
CC patient afflicted with ovarian cancer comprising providing to cells of  
CC the patient an antisense oligonucleotide complementary to a marker of the  
CC invention. The markers are useful for assessing if a patient is afflicted  
CC with ovarian cancer, which involves comparing the level of expression of  
CC a marker in a patient sample and a normal level of expression of the  
CC marker in a control non-ovarian cancer sample. A difference between the  
CC expression levels indicates ovarian cancer. The level of expression of a  
CC marker corresponds to a secreted protein or to a transcribed  
CC polynucleotide or its portion. The level of expression of the marker is  
CC assessed by detecting the presence in the sample, a protein or protein  
CC fragment corresponding to the marker. The presence of protein or protein  
CC fragment is detected using an antibody that specifically binds with the  
CC protein or protein fragment. Alternatively, the level of expression of  
CC the marker is assessed by detecting the presence of a transcribed  
CC polynucleotide which anneals with the marker or anneals with a portion of  
CC the polynucleotide comprising the marker, under stringent conditions. The  
CC marker is also used for monitoring the progression of ovarian cancer in a  
CC patient which involves detecting expression of the marker in a patient  
CC sample at a first point in time, repeating the method at a subsequent  
CC time and comparing the level of expression. The method is carried out  
CC using an ovarian tissue sample. A composition comprising a marker,  
CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
CC This sequence represents a human ovarian cancer DNA marker of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2285 BP; 749 A; 601 C; 523 G; 412 T; 0 U; 0 Other;  
Query Match 89.0%; Score 1139.4; DB 5; Length 2285;  
Best Local Similarity 92.7%; Pred. No. 1.6e-200;  
Matches 1248; Conservative 0; Mismatches 26; Indels 73; Gaps 2;  
QY 1 GAAAGATGGCTCC-CCGAGGAGGTACGGCTCTACTGCCACCTCTTCCAGCTCCACC 59  
DB 37 GAAAGATGGCTCCGGCGAAGGAGGTACGGCTCTACTGCCACCTCTTCCAGCTCCACC 96  
QY 60 GCCCGCGCAGCAGGGAAGCAAGGCGCTCGGGAGNTTCAGCCGTGAAGCAA 119  
DB 97 GCCCGCGCAGCAGGGAAGCAAGGCGCTCGGGAGNTTCAGCCGTGAAGCAA 156  
QY 120 GTGCAGATAGTGGCTTGTGGTATTAAAGATAATCAACATTATCAAGGAAGGACAA 179  
DB 157 GTGCAGATAGTGGCTTGTGGTATTAAAGATAATCAACATTATCAAGGAAGGACAA 216

QY 180 GGAACCTGAAGTTGTTCAAGAGTGTCTTTGGGTCCTGTTGTAGAAAGATCGGCTTGAAATT 239  
DB 217 GGAACCTGAAGTTGTTCAAGAGTGTCTTTGGGTCCTGTTGTAGAAAGATCGGCTTGAAATT 276  
QY 240 ACCAACTGCTTTCTCTTCCCTCAGCACACAGAGGATGATGCTGACATTTGATGAAG- 294  
DB 277 ACCAACTGCTTTCTCTTCCCTCAGCACACAGAGGATGATGCTGACATTTGATGAAGTTCA 336  
QY 295 ----- 294  
DB 337 TTCATAGTCAGCATACATCAGAACTTCGTTCTGTTATTGCTGAGTAGTATTCATTCA 396  
QY 295 -----TCCAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAAACATTGATCAT 347  
DB 397 AATGTAATCCAATATCAGATGGAATGATGCGGAGCCTTCGCCATGTAAACATTGATCAT 456  
QY 348 CTTACGCTGGGCTGGTATCAGTCCACATACTATGGCTCATTCGTTTACC CGGACACTCCGTG 407  
DB 457 CTTACGCTGGGCTGGTATCAGTCCACATACTATGGCTCATTCGTTTACC CGGACACTCCGTG 516  
QY 408 GACTCTCAGTTAGTTTACCGATGCGCAATGGAAGATCTGTCGTTCTCATTTATGATCCC 467  
DB 517 GACTCTCAGTTAGTTTACCGATGCGCAATGGAAGATCTGTCGTTCTCATTTATGATCCC 576  
QY 468 ATAAAAAATGCGCAAGGATCTCTCTCAGTAAAGGATACAGACTGACTCCTTAAACTGATG 527  
DB 577 ATAAAAAATGCGCAAGGATCTCTCTCAGTAAAGGATACAGACTGACTCCTTAAACTGATG 636  
QY 528 GAAAGTTGTTAAAGAAAAGGATTTTCCCTCTGAAGCATTTGAAAAAAGCAAAATATACCTTT 587  
DB 637 GAAAGTTGTTAAAGAAAAGGATTTTCCCTCTGAAGCATTTGAAAAAAGCAAAATATACCTTT 696  
QY 588 GAGTACATGTTTGAAGAAAGTCCGATTTGTAATTAATAAATTCACATCTCATGATGCTCTA 647  
DB 697 GAGTACATGTTTGAAGAAAGTCCGATTTGTAATTAATAAATTCACATCTCATGATGCTCTA 756  
QY 648 ATGTGCGGAACCTGAAAAAGAGTCAGCTCTGCAGATAAATGATGCTCAGCCTTGCC 707  
DB 757 ATGTGCGGAACCTGAAAAAGAGTCAGCTCTGCAGATAAATGATGCTCAGCCTTGCC 816  
QY 708 AGCAGCAATCATTTGGGGAAGAAATCTACAGTTGCTGATGGAAGAGTGGATGAAATGAGC 767  
DB 817 AGCAGCAATCATTTGGGGAAGAAATCTACAGTTGCTGATGGAAGAGTGGATGAAATGAGC 876  
QY 768 CAAGATATAGTTAAATCAACACATACATAGGAGTAATCTAGTAACACACAGCAGAGAA 827  
DB 877 CAAGATATAGTTAAATCAACACATACATAGGAGTAATCTAGTAACACACAGCAGAGAA 936  
QY 828 CATCAGTATCAGACGCTGCGCAGCAGAGAAATATGACGCCAGAGCCGAGGAGAACCC 887  
DB 937 CATCAGTATCAGACGCTGCGCAGCAGAGAAATATGACGCCAGAGCCGAGGAGAACCC 996  
QY 888 CCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCAAAACCCACACAGCGCGCTGCGCAGGATG 947  
DB 997 CCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCAAAACCCACACAGCGCGCTGCGCAGGATG 1056  
QY 948 GACTCTGCTGCTCATTTGAGGCGCAGATAAACAATTTACTGCCAGACATCTCAGAGGTTCACT 1007  
DB 1057 GACTCTGCTGCTCATTTGAGGCGCAGATAAACAATTTACTGCCAGACATCTCAGAGGTTCACT 1116  
QY 1008 GCCCAAACTTAGGCAAGCTCTTCATGCGCCAGGCTCTTCAAGAAATACAAACAATAAGAA 1067  
DB 1117 GCCCAAACTTAGGCAAGCTCTTCATGGCCAGGCTCTTCAAGAAATACAAACAATAAGAA 1176  
QY 1068 AAGGAAGTTTCCAGAAAAGAGTTAAATGAACTTTGAAAGTCAACACAGGCGCAACTCTT 1127  
DB 1177 AAGGAAGTTTCCAGAAAAGAGTTAAATGAACTTTGAAAGTCAACACAGGCGCAACTCTT 1236  
QY 1128 GGAAGAAATATATTGTCATATTGAAAAAGCAGAGGATTTCTTTAGTGTGCTATTGCCGATT 1187  
DB 1237 GGAAGAAATATATTGTCATATTGAAAAAGCAGAGGATTTCTTTAGTGTGCTATTGCCGATT 1296  
QY 1188 TTGCTATACAGTGTCTTTCTAGCCCATATAATAAATAAATAAATAAATAAATAAATAAATAA 1247

Db 1297 TTGGCTATAACAGTGTCTTTCTAGCCATAATAAATACACCTTGACACACAGCCA 1356  
Qy 1248 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1274  
Db 1357 CAGCAACACACAGCACCAACCAAAA 1383  
RESULT 8  
ID ADC30341 standard; cDNA; 2455 BP.  
XX AC ADC30341;  
XX DT 18-DEC-2003 (first entry)  
XX DE Human novel cDNA sequence, SEQ ID NO:423.  
XX KW Human; diagnostic; drug screening; forensics; gene mapping;  
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KW ulcers; osteoporosis; autoimmune disease; cancer;  
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;  
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KW gene therapy; chromosome 9p22; gene; ss.  
XX OS Homo sapiens.  
XX FN WO2003029271-A2.  
XX PD 10-APR-2003.  
XX PP 24-SEP-2002; 2002WO-US030474.  
XX PR 24-SEP-2001; 2001US-0324631P.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX WPI; 2003-371981/35.  
XX P-PSDB; ADC31312.  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
XX treating conditions such as neurodegenerative diseases, anemias, platelet  
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
XX cancer.  
XX Claim 1; SEQ ID NO 423; 1185pp; English.  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
XX invention also relates to nucleic acid sequences over 99% identical with  
XX the novel human cDNAs. The invention additionally encompasses expression  
XX vectors and host cells comprising a nucleic acid of the invention; the  
XX recombinant production of a polypeptide of the invention; an antibody  
XX against a polypeptide of the invention; a method of detecting  
XX polynucleotides or polypeptides of the invention; and methods of  
XX identifying a compound which binds to a polypeptide of the invention. The  
XX invention further discloses methods of preventing, treating or  
XX ameliorating a medical condition; kits comprising polynucleotide probes  
XX and/or monoclonal antibodies for carrying out the methods of the  
XX invention; methods for the identification of compounds that modulate the  
XX expression or activity of the polynucleotide and/or polypeptide; and 767  
XX contig sequences corresponding to the cDNA sequences of the invention  
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
XX -ADC33394). The nucleic acids and polypeptides of the invention are  
XX useful in diagnostics, drug screening, forensics, gene mapping, in the  
XX identification of mutations responsible for genetic disorders or other  
XX traits, for assessing biodiversity, and in producing many other types of

CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human cDNA sequence of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 2455 BP; 803 A; 542 C; 523 G; 587 T; 0 U; 0 Other;  
Query Match 86.5%; Score 1107; DB 10; Length 2455;  
Best Local Similarity 96.0%; Pred. No. 1.5e-194;  
Matches 1165; Conservative 0; Mismatches 0; Indels 48; Gaps 1;  
Qy 1 GAAAGATGGCGTCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Db 8 GAAAGATGGCGTCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 67  
Qy 61 CCGCGCAGCAGGAGGAAAGGCAAGGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
Db 68 CCGCGCAGCAGGAGGAAAGGCAAGGCGGCTCGGAGATTCAGCCGTGAAGCAAG 127  
Qy 121 TGCAGATAGATGGCCTT----- 137  
Db 128 TGCAGATAGATGGCCTTTTCTCTGTCAGGTATCTCTCAGAAGCCCTTCTCGATCCCC 187  
Qy 138 -----GTGGTATTAAAGATAATCAACATTATCAAGAAGAGGCAAGGAAGTCAAGTTG 192  
Db 188 CTAAGTGGTATTAAAGATAATCAACATTATCAAGAAGAGGCAAGGAAGTCAAGTTG 247  
Qy 193 TTCAAGGAGTGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAAATTTACCAATTCGCTTC 252  
Db 248 TTCAAGGAGTGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAAATTTACCAATTCGCTTC 307  
Qy 253 CTTTCCTCAGCAGCAGAGAGATGATGCTGACTTTGATGAAGTCCATATCAGATGAAA 312  
Db 308 CTTTCCTCAGCAGCAGAGAGATGATGCTGACTTTGATGAAGTCCATATCAGATGAAA 367  
Qy 313 TGATGGGAGCCTTGGGCATGTAAACATTGATCATCTTACGTGGGTGGTATCAGTCCA 372  
Db 368 TGATGGGAGCCTTGGGCATGTAAACATTGATCATCTTACGTGGGTGGTATCAGTCCA 427  
Qy 373 CATATATGGCTCATTTGGTTTACCGGGCACTCCTGGACTCTCAGTTTAGTTACAGCATG 432  
Db 428 CATATATGGCTCATTTGGTTTACCGGGCACTCCTGGACTCTCAGTTTAGTTACAGCATG 487  
Qy 433 CCATTGAAGAAATCTGTGGTTCTCATTTATGATCCATATAAAACTGCCCAAGGATCTCTCT 492  
Db 488 CCATTGAAGAAATCTGTGGTTCTCATTTATGATCCCAATAAAACTGCCCAAGGATCTCTCT 547  
Qy 493 CACTAAAGGCATACAGACTGACTCCTTAACTGATGAAGTTTGTAAAGAAAGGATTTTT 552  
Db 548 CACTAAAGGCATACAGACTGACTCCTTAACTGATGAAGTTTGTAAAGAAAGGATTTTT 607  
Qy 553 CCCTCTGAAGCATTTGAAAAAGCAAAATATCACTTTGAGTACATGTTTGAAGAAAGTCCGA 612  
Db 608 CCCTCTGAAGCATTTGAAAAAGCAAAATATCACTTTGAGTACATGTTTGAAGAAAGTCCGA 667  
Qy 613 TTGTAATTAATAATTCATCTCATCTGATCAATGCTCTTAATGTGGGAACCTTGAAAGAGTCAG 672  
Db 668 TTGTAATTAATAATTCATCTCATCTGATCAATGCTCTTAATGTGGGAACCTTGAAAGAGTCAG 727  
Qy 673 CTGTTGCAGATAAACATGAATTCCTCAGCCTTCCAGCAGCAATCATTTGCGGAAGATC 732  
Db 728 CTGTTGCAGATAAACATGAATTCCTCAGCCTTCCAGCAGCAATCATTTGCGGAAGATC 787  
Qy 733 TACAGTTGCTGATGGAGCAGAGTGGATGAATAGCCAAATAGTAAATATACAAACAT 792

Db 788 TACAGTTGCTGATGAGCAGAGTGATGAAATGAGCCAGATATAGTTAAATACACACAT 847  
 QY 793 ACATGAGGATAGTAAACACACAGCAGACAGAAACATCAGTATCAGCAGCGTCCCGAGC 852  
 Db 848 ACATGAGGATAGTAAACACACAGCAGCAGAAACATCAGTATCAGCAGCGTCCCGAGC 907  
 QY 853 AGGAGAAATGACAGCCAGAGCCAGAGGAGAAACCCCGCTCCCTCAGGAGGACCTGTCCA 912  
 Db 908 AGGAGAAATGACAGCCAGAGCCAGAGGAGAAACCCCGCTCCCTCAGGAGGACCTGTCCA 967  
 QY 913 AACTCTTCAACACACAGCGCCCTGCAGGATGGAATCGCTGCTATTGCGAGCCAGA 972  
 Db 968 AACTCTTCAACACACAGCGCCCTGCAGGATGGAATCGCTGCTATTGCGAGCCAGA 1027  
 QY 973 TAAACACTTACTGCCAGAACATCAAGGAGTTCACTGCCCAAACTTAGGCAAGCTCTTCA 1032  
 Db 1028 TAAACACTTACTGCCAGAACATCAAGGAGTTCACTGCCCAAACTTAGGCAAGCTCTTCA 1087  
 QY 1033 TGGCCAGGCTCTTCAAGAAATACAACTAAGAAAAGGAAAGTTTCCAGAAAAGAGTTA 1092  
 Db 1088 TGGCCAGGCTCTTCAAGAAATACAACTAAGAAAAGGAAAGTTTCCAGAAAAGAGTTA 1147  
 QY 1093 ACATGAACCTTGAAGTACACAGGAGCACTCTTGGAAAGAAATATATTGCAATATTGAA 1152  
 Db 1148 ACATGAACCTTGAAGTACACAGGAGCACTCTTGGAAAGAAATATATTGCAATATTGAA 1207  
 QY 1153 AAGCACAGAGAT 1165  
 Db 1208 AAGCACAGAGAT 1220

RESULT 9  
 ID ADE29036  
 XX AC ADE29036; standard; DNA; 1045 bp.  
 XX AC ADE29036;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human EIF3S3 selected interacting domain DNA - SEQ ID 22.  
 XX KW anti-HIV; SID; selected interacting domain; HIV; gene therapy; siRNA;  
 KW silencing RNA; HIV-1 integrase; LEDGF; MCM7; HBO1; Snurportin; VBP1;  
 KW Transportin-SR; EIF3S3; ds; gene; human.  
 XX OS Homo sapiens.  
 XX PN WO2003046176-A2.  
 XX PD 05-JUN-2003.  
 XX PF 26-NOV-2002; 2002MO-EP013868.  
 XX PR 26-NOV-2001; 2001US-0333346P.  
 XX PR 31-MAY-2002; 2002US-0385132P.  
 XX PA (HYBR-) HYBRIGENICS.  
 XX PI Legrain P, Rain J, Benarous R, Emiliani S, Berlioz-Torrent C;  
 PI Blot G;  
 XX WPI; 2003-505199/47.  
 XX DR P-PSDB; ADE29059.  
 XX PT New complex between two interacting proteins, useful for screening  
 PT molecules that inhibit human immunodeficiency virus or for preparing a  
 PT medicament for treating HIV-1.  
 XX PS Claim 10; SEQ ID NO 22; 102pp; English.  
 XX CC The invention relates to a novel complex between two interacting proteins  
 CC listed within the specification. The complex of the invention  
 CC demonstrates anti-HIV activity whilst the SID (selected interacting

CC domains) and polypeptides may be useful for screening molecules that  
 CC inhibit human immunodeficiency virus (HIV), as well as during gene  
 CC therapy procedures. The siRNAs (silencing RNAs) targeted against cellular  
 CC proteins interacting with HIV-1 integrase, LEDGF, MCM7, HBO1, Snurportin,  
 CC VBP1, Transportin-SR and EIF3S3, are useful in preparing a medicament for  
 CC treating HIV-1. The current sequence is that of the human SID DNA of the  
 CC invention.  
 XX SQ Sequence 1045 BP; 326 A; 243 C; 237 G; 239 T; 0 U; 0 Other;  
 Query Match 81.6%; Score 1045; DB 10; Length 1045;  
 Best Local Similarity 100.0%; Fred. No. 3.6e-183;  
 Matches 1045; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 20 GGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGGCGAGGAAAGG 79  
 Db 1 GGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGGCGAGGAAAGG 60  
 QY 80 CAAAGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAGTGCAGATAGATGGCTTGT 139  
 Db 61 CAAAGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAGTGCAGATAGATGGCTTGT 120  
 QY 140 GGTATTAAAGATAAATCAAAACATTATCAAGAAAGCAAGCAAGTGAAGTTGTTCAAG 199  
 Db 121 GGTATTAAAGATAAATCAAAACATTATCAAGAAAGCAAGCAAGTGAAGTTGTTCAAG 180  
 QY 200 AGTGCTTTGGGCTCTGGTGTGAAGATCGGCTTGAAATTTACCAACTGCTTTCTTTCCC 259  
 Db 181 AGTGCTTTGGGCTCTGGTGTGAAGATCGGCTTGAAATTTACCAACTGCTTTCTTTCCC 240  
 QY 260 TCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATGATCG 319  
 Db 241 TCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATGATCG 300  
 QY 320 GAGCCTTCGCCATGTAAACATTTGATCATCTTCACTGGGCTGGTATCAGTCCACATCTA 379  
 Db 301 GAGCCTTCGCCATGTAAACATTTGATCATCTTCACTGGGCTGGTATCAGTCCACATCTA 360  
 QY 380 TGGCTCATTTGTTACCGGGCACTCTCGGACTCTCAGTTTACGATGTCATTTGA 439  
 Db 361 TGGCTCATTTGTTACCGGGCACTCTCGGACTCTCAGTTTACGATGTCATTTGA 420  
 QY 440 AGAATCTGCTCTCTCTATTTATGATCCATAAACTGCCCAAGGATCTCTCTCACTAAA 499  
 Db 421 AGAATCTGCTCTCTCTATTTATGATCCATAAACTGCCCAAGGATCTCTCTCACTAAA 480  
 QY 500 GGCATACAGACTGACTCTCTAACTGATGGAAAGTTTGTAAAGAAAGGATTTTCCCCTGA 559  
 Db 481 GGCATACAGACTGACTCTCTAACTGATGGAAAGTTTGTAAAGAAAGGATTTTCCCCTGA 540  
 QY 560 AGCATTGAAAAGCAAAATATCACCCTTTGATGATCATCTTTGAAGAGTCCGATTTGTA 619  
 Db 541 AGCATTGAAAAGCAAAATATCACCCTTTGATGATCATCTTTGAAGAGTCCGATTTGTA 600  
 QY 620 TAAAAATTCAATCTGATCAATGTCTTAATGTGGAACTTTGAAAGAGTCAAGCTGTTGC 679  
 Db 601 TAAAAATTCAATCTGATCAATGTCTTAATGTGGAACTTTGAAAGAGTCAAGCTGTTGC 660  
 QY 680 AGATAAACATGAATGCTGACGCTTGCAGCAGCAATCATTTCGGGGAAGATCTACAGTT 739  
 Db 661 AGATAAACATGAATGCTGACGCTTGCAGCAGCAATCATTTCGGGGAAGATCTACAGTT 720  
 QY 740 GCTGATGACAGAGTGGATGAAATGAGCCAAAGATATAGTTAAATACACATACATGAG 799  
 Db 721 GCTGATGACAGAGTGGATGAAATGAGCCAAAGATATAGTTAAATACACATACATGAG 780  
 QY 800 GAATACCTAGTAAACACAGCAGCAAGAAACATCAGTATCAGCAGCGTCCGCGAGGAGAA 859  
 Db 781 GAATACCTAGTAAACACAGCAGCAAGAAACATCAGTATCAGCAGCGTCCGCGAGGAGAA 840  
 QY 860 TATGCGCGCCAGCGGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTCCAACTCTT 919  
 Db 841 TATGCGCGCCAGCGGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTGTCCAACTCTT 900

XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;  
KW immunogenic; vaccine; ss.  
XX  
OS Homo sapiens.  
XX

FN WO200151513-A2.

XX 19-JUL-2001.

XX 16-JAN-2001; 2001WO-US001575.

XX 14-JAN-2000; 2000US-0176722P.

XX (CORI-) CORIXA CORP.

XX Algate PA;

XX WPI; 2001-425866/45.

XX Novel ovarian tumor proteins, and nucleic acids encoding them, used to  
PT treat and diagnose cancers, particularly ovarian cancer.  
XX

PS Claim 5; Page 317; 338pp; English.

XX AAH82377 to AAH83878 represent human ovarian tumour-associated  
CC polynucleotide sequences which encode ovarian tumour proteins. The  
CC ovarian tumour protein and polynucleotide sequences have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. The  
CC ovarian tumour proteins and polynucleotides can be used to inhibit the  
CC development of cancer, particularly ovarian cancer. They can also be used  
CC to diagnose the onset and progression of cancer  
XX

SQ Sequence 793 BP; 220 A; 176 C; 185 G; 197 T; 0 U; 15 Other;

Query Match 43.9%; Score 562.4; DB 5; Length 793;  
Best Local Similarity 97.6%; Pred. No. 2.7e-94;  
Matches 603; Conservative 0; Mismatches 11; Indels 4; Gaps 3;

Qy 52 GCTCCACCGCGCGCAGCAGGAAAGCAAGGCGGCTCGGGAGATTCAGCCG 111

Db 11 GCTCCACCGCGCGCAGGAAAGCAAGGCGGCTCGGGAGATTCAGCCG 70

Qy 112 TGAAGCAAGTCAGATAGATGCGCCTTGTTGTTAATCAAAATATCAAGAAG 171

Db 71 TGAAGCAAGTCAGATAGATGCGCCTTGTTGTTAATCAAAATATCAAGAAG 130

Qy 172 AAGGACAAGGAACTGAAGTTGTTCAAGAGTGTCTTTGGGTCTGGTTGTAGAAGTCGGC 231

Db 131 AAGGACAAGGAACTGAAGTTGTTCAAGAGTGTCTTTGGGTCTGGTTGTAGAAGTCGGC 190

Qy 232 TTGAATTTACCACTGCTTCTTCCCTCAGCACACAGAGGATGCTGACTTTGATG 291

Db 191 TTGAATTTACCACTGCTTCTTCCCTCAGCACACAGAGGATGCTGACTTTGATG 250

Qy 292 AAGTCCAATATCAGATGGAATGATGCGGAGCCTTCGCATGTAAACATTGATCATCTTC 351

Db 251 AAGTCCAATATCAGATGGAATGATGCGGAGCCTTCGCATGTAAACATTGATCATCTTC 310

Qy 352 ACGTGGGCTGGTATCAGTCCACATACATGCTCATTCGTTTACCCGGGCACTCCTGACT 411

Db 311 ACGTGGGCTGGTATCAGTCCACATACATGCTCATTCGTTTACCCGGGCACTCCTGACT 370

Qy 412 CTCAGTTTGTAGTACAGATGCCATTGAGAACTCTGCTCTCATTTATGATCCCATAA 471

Db 371 CTCAGTTTGTAGTACAGATGCCATTGAGAACTCTGCTCTCATTTATGATCCCATAA 430

Qy 472 AAACGTCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAG 531

Db 431 AAACGTCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAG 490

Qy 532 TTTGTAAAGAAAGGATTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGT 591

Db 491 TTTGTAAAGAAAGGATTTTCCCTGAA-CATTGAAAAAGCAAAATATCACCTTTGAGT 549  
Qy 592 ACATGTTTGAAGAGTG-CGATTGTAATTAAAAATTACATCT--GATCAATGTCCTAA 648  
Db 550 ACAIGTTTGAAGAGAGTGCCCGATTGTAATTAAAAATTACATCTTTGATCAATGTCCTAA 609  
Qy 649 TGTGGGAACCTTGAAAAAGA 666  
Db 610 TGTGGGAACCTTGAAAAA 627

Search completed: January 5, 2005, 13:00:20  
Job time : 689 secs

**This Page Blank (uspto)**

Qy 920 CAAACCAACACAGCGCCTGCCAGGATGGACTCGCTGCTCATTTGACGCCAGATAAAACAC 979  
Db 901 CAAACCAACACAGCGCCTGCCAGGATGGACTCGCTGCTCATTTGACGCCAGATAAAACAC 960  
Qy 980 TTACTGCCAGAACATCAAGAGATTCACTGCCCAAACTTAGGCAAGCTCTTCATGGGCCCA 1039  
Db 961 TTACTGCCAGAACATCAAGAGATTCACTGCCCAAACTTAGGCAAGCTCTTCATGGGCCCA 1020  
Qy 1040 GGCTCTTCAAGAAATACAACTAA 1064  
Db 1021 GGCTCTTCAAGAAATACAACTAA 1045

RESULT 10  
ADE29037  
ID ADE29037 standard; DNA; 908 BP.  
XX  
AC ADE29037;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human EIF3S3 selected interacting domain DNA - SEQ ID 23.  
XX  
KW anti-HIV; SID; selected interacting domain; HIV; gene therapy; siRNA;  
KW silencing RNA; HIV-1 integrase; LEDGF; MCM7; HB01; Snurportin; VBp1;  
KW Transportin-SR; EIF3S3; ds; gene; human.  
XX  
XX Homo sapiens.  
XX  
XX WO2003046176-A2.  
XX  
XX 05-JUN-2003.  
XX  
XX 26-NOV-2002; 2002WO-EP013868.  
XX  
XX 26-NOV-2001; 2001US-0333346P.  
XX  
XX 31-MAY-2002; 2002US-0385132P.  
XX  
XX (HYBR-) HYBRIGENICS.  
XX  
XX Legrain P, Rain J, Benarous R, Emiliani S, Berlioz-Torrent C;  
XX Blot G;  
XX  
XX WPI; 2003-505199/47.  
XX  
XX P-PSDB; ADE29060.  
XX  
XX New complex between two interacting proteins, useful for screening  
XX molecules that inhibit human immunodeficiency virus or for preparing a  
XX medicament for treating HIV-1.  
XX  
XX Claim 10; SEQ ID NO 23; 102pp; English.

XX The invention relates to a novel complex between two interacting proteins  
XX listed within the specification. The complex of the invention  
XX demonstrates anti-HIV activity whilst the SID (selected interacting  
XX domains) and polypeptides may be useful for screening molecules that  
XX inhibit human immunodeficiency virus (HIV), as well as during gene  
XX therapy procedures. The siRNAs (silencing RNAs) targeted against cellular  
XX proteins interacting with HIV-1 integrase, LEDGF, MCM7, HB01, Snurportin,  
XX VBp1, Transportin-SR and EIF3S3, are useful in preparing a medicament for  
XX treating HIV-1. The current sequence is that of the human SID DNA of the  
XX invention.

XX  
SQ Sequence 908 BP; 285 A; 200 C; 210 G; 213 T; 0 U; 0 Other;  
Query Match 70.9%; Score 908; DB 10; Length 908;  
Best Local Similarity 100.0%; Pred. No. 6.1e-158;  
Matches 908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 17 CAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTTCCACCGCGCGCAGCGGAA 76  
Db 1 CAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTTCCACCGCGCGCAGCGGAA 60

Qy 77 AGCAAGGCAAGGCGGCTCGGAGATTTCAGCGGTGAAGCAAGTGCAGATAGATGGCT 136  
Db 61 AGCAAGGCAAGGCGGCTCGGAGATTTCAGCGGTGAAGCAAGTGCAGATAGATGGCT 120  
Qy 137 TGTGGTATTAAGATAATCAAAATATCAAGAAAGGAGCAAGGAACTGAAGTTGTTC 196  
Db 121 TGTGGTATTAAGATAATCAAAATATCAAGAAAGGAGCAAGGAACTGAAGTTGTTC 180  
Qy 197 AGGAGTCTTTGGGCTCTGGTTGTAGAAGATCGGCTTGAATTAACCAACTGCTTTCTTT 256  
Db 181 AGGAGTCTTTGGGCTCTGGTTGTAGAAGATCGGCTTGAATTAACCAACTGCTTTCTTT 240  
Qy 257 CCCTCAGCACACAGAGGATGATGCTGATCTTTGATGAAGTCCAATATCAGATGGAATGAT 316  
Db 241 CCCTCAGCACACAGAGGATGATGCTGATCTTTGATGAAGTCCAATATCAGATGGAATGAT 300  
Qy 317 GCGGAGCCTTCGCCATGTAACATTTGATCATCTTTACGTGGGCTGGTATCAGTCCACATA 376  
Db 301 GCGGAGCCTTCGCCATGTAACATTTGATCATCTTTACGTGGGCTGGTATCAGTCCACATA 360  
Qy 377 CTATGGCTCATCTGTTACCGGCGCACCTCGGACTCTCAGTTTATACCAAGCATGCCAT 436  
Db 361 CTATGGCTCATCTGTTACCGGCGCACCTCGGACTCTCAGTTTATACCAAGCATGCCAT 420  
Qy 437 TGAAGAAATCTGCTCTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCACT 496  
Db 421 TGAAGAAATCTGCTCTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCACT 480  
Qy 497 AAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGAAGAAAGGATTTTTCCTCC 556  
Db 481 AAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGAAGAAAGGATTTTTCCTCC 540  
Qy 557 TGAAGCAATGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAAAGTGCAGATTGT 616  
Db 541 TGAAGCAATGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAAAGTGCAGATTGT 600  
Qy 617 AATTAAAAATTCACATCTGATCAATGTCTTAATGTGGAACTTGAAGAAAGTGCAGCTGT 676  
Db 601 AATTAAAAATTCACATCTGATCAATGTCTTAATGTGGAACTTGAAGAAAGTGCAGCTGT 660  
Qy 677 TGCAGATAAATCATGAATTTGCTCAGCCTTGCAGCAGCAATCATTTGGGAAAGAAATCTTACA 736  
Db 661 TGCAGATAAATCATGAATTTGCTCAGCCTTGCAGCAGCAATCATTTGGGAAAGAAATCTTACA 720  
Qy 737 GTTGTCTGATGACAGAGTGGATGAAATGAGCCAGATATAGTTAAATACACATACAT 796  
Db 721 GTTGTCTGATGACAGAGTGGATGAAATGAGCCAGATATAGTTAAATACACATACAT 780  
Qy 797 GAGGAATACCTAGTAAACCAACAGCAGCAAAACATCAGTATCAGCAGCGCTCGCCAGCAGGA 856  
Db 781 GAGGAATACCTAGTAAACCAACAGCAGCAAAACATCAGTATCAGCAGCGCTCGCCAGCAGGA 840  
Qy 857 GAATATCAGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAGGAGGACCTGTCCAAACT 916  
Db 841 GAATATCAGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAGGAGGACCTGTCCAAACT 900  
Qy 917 CTTCAAAAC 924  
Db 901 CTTCAAAAC 908  
RESULT 11  
ABQ55406  
ID ABQ55406 standard; cDNA; 774 BP.  
XX  
XX AC ABQ55406;  
XX  
XX 22-AUG-2002 (first entry)  
XX  
XX Human ovarian antigen HNOBE83 cDNA, SEQ ID NO:1286.  
XX  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;



ovarian cancer; breast cancer; tumour; reproductive system disorder; infertility; pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection; inflammatory condition; immune disorder; blood disorder; cardiovascular disorder; respiratory disorder; neurological disorder; gastrointestinal disorder; urinary system disorder; drug screening; gene therapy; chromosome mapping; forensic analysis; antibody preparation; cytostatic; immunomodulatory; neuroprotective; antiinflammatory; gynaecological; reproductive; gene; ss.

Homo sapiens.  
WO200200677-A1.  
03-JAN-2002.  
07-JUN-2001; 2001WO-US018569.  
07-JUN-2000; 2000US-0209467P.  
(HUMA-) HUMAN GENOME SCI INC.  
Birse CE, Rosen CA;  
WPI; 2002-147878/19.  
P-PSDB; ABP42329.

Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases.

Claim 1; SEQ ID NO 1286; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents cDNA encoding a human ovarian antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 774 BP; 236 A; 162 C; 180 G; 193 T; 0 U; 3 Other;

Query Match 56.5%; Score 723.6; DB 6; Length 774;  
Best Local Similarity 97.4%; Pred. No. 5.5e-124;  
Matches 755; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

15 CGCAAGGAGGTACCGGCTCTACTGCCACTCTTCCAGCTCCACCGCGCGCAGCAGGG 74  
2 CGCAAGGAGGTACCGGCTCTACTGCCACTCTTCCAGCTCCACCGCGCGCAGCAGGG 61

QY 75 AAAGGCAAGGCAAGGCGCGCTCGGAGATTTCAGCCGTGAAGCAAGTCGACAGATAGATGGC 134  
DB 62 AAAGGCAAGGCAAGGCGCGCTCGGAGATTTCAGCCGTGAAGCAAGTCGACAGATAGATGGC 121  
QY 135 CTTGTGGTATTAAAGATATCAAACTTATCAAGAAAGGACAAGGAACCTGAAGTTGTT 194  
DB 122 CTTGTGGTATTAAAGATATCAAACTTATCAAGAAAGGACAAGGAACCTGAAGTTGTT 181  
QY 195 CAAGGAGTGTCTTTGGGCTCTGGTTGTAGAGAGATCGGCTTGAATATCAACCACTGCTTCTCT 254  
DB 182 CAAGGAGTGTCTTTGGGCTCTGGTTGTAGAGAGATCGGCTTGAATATCAACCACTGCTTCTCT 241  
QY 255 TTCCCTCAGACACAGAGGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATG 314  
DB 242 TTCCCTCAGACACAGAGGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATG 301  
QY 315 ATGCGGAGCCTTCGCCATGTAAACATTTGATCACTTTCAGTGGGCTGGTATCAGTCCACA 374  
DB 302 ATGCGGAGCCTTCGCCATGTAAACATTTGATCACTTTCAGTGGGCTGGTATCAGTCCACA 361  
QY 375 TACTATGGCTCATTCGTTTACCCGGGCACTCTCTGGACTCTCAGTTTAGTTTACCAGCATGCC 434  
DB 362 TACTATGGCTCATTCGTTTACCCGGGCACTCTCTGGACTCTCAGTTTAGTTTACCAGCATGCC 421  
QY 435 ATTGAAGATCTGTCGTTCTCATTTATGATCCATAAAACTGCCAAGGATCTCTCTCA 494  
DB 422 ATTGAAGATCTGTCGTTCTCATTTATGATCCATAAAACTGCCAAGGATCTCTCTCA 481  
QY 495 CTAAAGGCATACAGACTGACTCTCTAAACTGATGGAGTTTCTAAGAAAGGATTTTTC 554  
DB 482 CTAAAGGCATACAGACTGACTCTCTAAACTGATGGAGTTTCTAAGAAAGGATTTTTC 541  
QY 555 CCTGAAGCATTAAGCAAAATATCACTTTCAGTACATGTTTGAAGAAAGTG-CCGAT 613  
DB 542 CCTGAAGCATTAAGCAAAATATCACTTTCAGTACATGTTTGAAGAAAGTGCCCGAT 601  
QY 614 TGTAAATTAAGAAATTCACATCTGATCAATGTCCTAATGTGGAACTTGAAGAAAGTCAAGC 673  
DB 602 TGTAAATTAAGAAATTCACATCTGATCAATGTCCTAATGTGGAACTTGAAGAAAGTCAAGC 661  
QY 674 TGTTCAGATAAAACATGAATTCGCTCAGCCTTTCAGCAGCAGCAATCAATTTGGGGAAAGATCT 733  
DB 662 TGTTCAGATAAAACATGAATTCGCTCAGCCTTTCAGCAGCAGCAATCAATTTGGGGAAAGATCT 719  
QY 734 ACAGTTCTGATGACAGAGTGGATGAAATGAGCCCAAGATATAGTTAAATACAAC 788  
DB 720 ACNGTTGCTGATGGACCAAAATGGATTAATGAGCCAGGAAATAGTAAATNCCAC 774

RESULT 12

AAS25348/c

ID AAS25348 standard; cDNA; 621 BP.

XX AAS25348;

AC AAS25348;

XX 07-NOV-2001 (first entry)

DT 07-NOV-2001 (first entry)

DE Human ovarian PCR-subtracted cDNA library clone #1433.

XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;  
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;  
KW primer; probe.  
XX Homo sapiens.

OS Homo sapiens.

XX WO200157207-A2.

PN WO200157207-A2.

XX 09-AUG-2001.

PD 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US003733.

PF 05-FEB-2001; 2001WO-US003733.

XX 04-FEB-2000; 2000US-0180403P.

PR 04-FEB-2000; 2000US-0180403P.





Db 181 GAACTGAAGTTGTTTCAAGGAGTGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAATA 240  
QY CCAACTGCTTTCCTTCCCTCAGCACAGAGGATGATGCTGACTTTTGATGAAGTCCAAT 300  
Db 241 CCAACTGCTTTCCTTCCCTCAGCACAGAGGATGATGCTGACTTTTGATGAAGTCCAAT 300  
QY 301 ATCAGATGGAATGATGCGGAGCCTTGCCTATGTAAACATTTGATCATCTTCACTGGGCT 360  
Db 301 ATCAGATGGAATGATGCGGAGCCTTGCCTATGTAAACATTTGATCATCTTCACTGGGCT 360  
QY 361 GGTATCAGTCACATATATGCTGCTCATTCCTTCACTTTATGATCCCAAAAAGTGGCC 420  
Db 361 GGTATCAGTCACATATATGCTGCTCATTCCTTCACTTTATGATCCCAAAAAGTGGCC 420  
QY 421 GTTACCAGATGCAATTTGAAGATCTGCTTCTTCACTTTATGATCCCAAAAAGTGGCC 480  
Db 421 GTTACCAGATGCAATTTGAAGATCTGCTTCTTCACTTTATGATCCCAAAAAGTGGCC 480  
QY 481 AAGGATCTCTCTCACTAAAGCATACAGATGCTCTTAACTGATGGAAGTTTGTAAAG 540  
Db 481 AAGGATCTCTCTCACTAAAGCATACAGATGCTCTTAACTGATGGAAGTTTGTAAAG 540  
QY 541 AAAAGGATTTTCCCTGGAAGCATTTGAAAAAGCAATATCACTTTGATGATGTTG 600  
Db 541 AAAAGGATTTTCCCTGGAAGCATTTGAAAAAGCAATATCACTTTGATGATGTTG 600  
QY 601 AAGAGTGGCGATTTGTAATTAATAATTCATCTGATCAATGCTTAACTGGAAGTTG 660  
Db 601 AAGAGTGGCGATTTGTAATTAATAATTCATCTGATCAATGCTTAACTGGAAGTTG 660  
QY 661 AAAAGAGTCACTGTTTGCAGATAAATGATGCTTCACTGCTGCGAGCAGCAATCAIT 720  
Db 661 AAAAGAGTCACTGTTTGCAGATAAATGATGCTTCACTGCTGCGAGCAGCAATCAIT 720  
QY 721 TGGGAAGATCTACAGTTCTGATGAGACAGATGATGAAATGAGCAGCAATGATTA 780  
Db 721 TGGGAAGATCTACAGTTCTGATGAGACAGATGATGAAATGAGCAGCAATGATTA 780  
QY 781 AATACACACATACATGAGGAACTAGTAAACAAAGCAGCAGCAATCACTATCAGC 840  
Db 781 AATACACACATACATGAGGAACTAGTAAACAAAGCAGCAGCAATCACTATCAGC 840  
QY 841 AGCGTCCAGCAGGAGAAATATCAGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 900  
Db 841 AGCGTCCAGCAGGAGAAATATCAGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 900  
QY 901 AGGACCTGTCCAACTCTTCAACACCACAGCGCTGCGAGATGGAATGGAATGGAAT 960  
Db 901 AGGACCTGTCCAACTCTTCAACACCACAGCGCTGCGAGATGGAATGGAATGGAAT 960  
QY 961 TTGCAGGCCAGATAAACAATCTACTGCGAGAAATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Db 961 TTGCAGGCCAGATAAACAATCTACTGCGAGAAATCAAGGAGTTCACTGCCCAAACTTAG 1020  
QY 1021 GCAAGCTCTTATGCGCCAGGCTCTTCAAGAAATCAACAACCTAAGAAAAAGGAGTTTCA 1080  
Db 1021 GCAAGCTCTTATGCGCCAGGCTCTTCAAGAAATCAACAACCTAAGAAAAAGGAGTTTCA 1080  
QY 1081 GAAAGAGTTTACATGAACTCTTGAAGTCACACAGGCGCACTCTTGAAGAAATATAT 1140  
Db 1081 GAAAGAGTTTACATGAACTCTTGAAGTCACACAGGCGCACTCTTGAAGAAATATAT 1140  
QY 1141 TTGCATATTGAAAGCAGCAGAGATTTCTTTAGTGTCAATTTGGCTATTAACAG 1200  
Db 1141 TTGCATATTGAAAGCAGCAGAGATTTCTTTAGTGTCAATTTGGCTATTAACAG 1200  
QY 1201 TGCTTTCTAGCCATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
Db 1201 TGCTTTCTAGCCATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
QY 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
Db 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

RESULT 2  
AX358735  
LOCUS AX358735  
DEFINITION Sequence 1 from Patent WO0192307.  
ACCESSION AX358735  
VERSION AX358735.1  
KEYWORDS GI:18675269  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Nicolette, C.A.  
TITLE Therapeutic compounds for ovarian cancer  
JOURNAL Patent: WO 0192307-A 1 06-DEC-2001;  
GENZYME CORPORATION (US)  
FEATURES  
source  
1..1280  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1280; DB 6; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 5e-264;  
Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCGACCTCTTCCAGCTCCACCG 60  
Db 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCGACCTCTTCCAGCTCCACCG 60  
QY 61 CCGGCGCAGCAGGAAAGCAAGCAAGGCGGCTCGGAGATTGAGCGCTGGAAGCAAG 120  
Db 61 CCGGCGCAGCAGGAAAGCAAGCAAGGCGGCTCGGAGATTGAGCGCTGGAAGCAAG 120  
QY 121 TCAGATAGATGGCTTGTGGTATTAAAGATTAATCAAACTTATCAAGAAAGGACCAAG 180  
Db 121 TCAGATAGATGGCTTGTGGTATTAAAGATTAATCAAACTTATCAAGAAAGGACCAAG 180  
QY 181 GAACTGAAGTTGTTCAAGGAGTGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAATTA 240  
Db 181 GAACTGAAGTTGTTCAAGGAGTGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAATTA 240  
QY 241 CCAACTGCTTTCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
Db 241 CCAACTGCTTTCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
QY 301 ATCAGATGGAATGATGCGGAGCCTTGCCTATGTAAACATTTGATCATCTTCACTGGGCT 360  
Db 301 ATCAGATGGAATGATGCGGAGCCTTGCCTATGTAAACATTTGATCATCTTCACTGGGCT 360  
QY 361 GGTATCAGTCACATATGCTGCTCATTCCTTCACTTTATGATCCCAAAAAGTGGCC 420  
Db 361 GGTATCAGTCACATATGCTGCTCATTCCTTCACTTTATGATCCCAAAAAGTGGCC 420  
QY 421 GTTACCAGATGCAATTTGAAGATCTGCTTCTTCACTTTATGATCCCAAAAAGTGGCC 480  
Db 421 GTTACCAGATGCAATTTGAAGATCTGCTTCTTCACTTTATGATCCCAAAAAGTGGCC 480  
QY 481 AAGGATCTCTCTCACTAAAGCATACAGATGCTCTTAACTGATGGAAGTTTGTAAAG 540  
Db 481 AAGGATCTCTCTCACTAAAGCATACAGATGCTCTTAACTGATGGAAGTTTGTAAAG 540  
QY 541 AAAAGGATTTTCCCTGGAAGCATTTGAAAAAGCAATATCACTTTGATGATGTTG 600  
Db 541 AAAAGGATTTTCCCTGGAAGCATTTGAAAAAGCAATATCACTTTGATGATGTTG 600  
QY 601 AAGAGTGGCGATTTGTAATTAATAATTCATCTGATCAATGCTTAACTGGAAGTTG 660  
Db 601 AAGAGTGGCGATTTGTAATTAATAATTCATCTGATCAATGCTTAACTGGAAGTTG 660

Qy 661 AAAAGAAGTCAGCTGTTCCAGATAAAACATGAATTGCTCAGCGTTGCCAGCAGCAATCAAT 720  
Db 661 AAAAGAAGTCAGCTGTTCCAGATAAAACATGAATTGCTCAGCGTTGCCAGCAGCAATCAAT 720  
Qy 721 TGGGGAGAAGTACAGTGTCTGAGGACAGAGTGGATGAATGAGCCAGATATAGTTA 780  
Db 721 TGGGGAGAAGTACAGTGTCTGAGGACAGAGTGGATGAATGAGCCAGATATAGTTA 780  
Qy 781 AATACAAACACATACATAGAGGAATATCTAGTAAACACAGCAGCAGAAAAATCAGTATCAGC 840  
Db 781 AATACAAACACATACATAGAGGAATATCTAGTAAACACAGCAGCAGAAAAATCAGTATCAGC 840  
Qy 841 AGCGTCGCAGCAGAGGAATATGACGCCAGAGCCGAGAGGAGAACCCCGCTCCCTGAGG 900  
Db 841 AGCGTCGCAGCAGAGGAATATGACGCCAGAGCCGAGAGGAGAACCCCGCTCCCTGAGG 900  
Qy 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCCGCTGCCAGGATGACCTCGCTGCTCA 960  
Db 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCCGCTGCCAGGATGACCTCGCTGCTCA 960  
Qy 961 TTGCAGGCCAGATAAAACATCTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Db 961 TTGCAGGCCAGATAAAACATCTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Qy 1021 GCAAGCTCTTCATGCCCCAGGCTCTTCAAGAAATACAACTAAGAAAGGAAGTTTCCA 1080  
Db 1021 GCAAGCTCTTCATGCCCCAGGCTCTTCAAGAAATACAACTAAGAAAGGAAGTTTCCA 1080  
Qy 1081 GAAAGAAGTTAAACATGAACCTTTGAAGTCAACCCAGCCGCACTTTGGAGAAGATATAT 1140  
Db 1081 GAAAGAAGTTAAACATGAACCTTTGAAGTCAACCCAGCCGCACTTTGGAGAAGATATAT 1140  
Qy 1141 TTGCATATTTGAAAGCAGCAGAGGATTTCTTTAGTGTCTATTCGCGATTTTGGCTATAACAG 1200  
Db 1141 TTGCATATTTGAAAGCAGCAGAGGATTTCTTTAGTGTCTATTCGCGATTTTGGCTATAACAG 1200  
Qy 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
Db 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
Qy 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
Db 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

RESULT 3  
AX779905 AX779905 1280 bp DNA linear PAT 14-JUL-2003  
LOCUS Sequence 2062 from Patent WO03039443.  
DEFINITION AX779905  
ACCESSION AX779905  
VERSION AX779905.1 GI:32696899  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Haerlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 2062 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE) ;  
Ludwig-Maximilian-Universitaet Muenchen (DE) ;  
PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)  
FEATURES  
source  
1. .1280  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 1280; DB 6; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 5e-264;

Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTTGCACCTCTTCCAGCTCCACCG 60  
Db 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Qy 61 CCGCGCAGCAGGGAAGGCAAGCGCGCTCGGGAGATTTCAGCCGCTGAAGCAAG 120  
Db 61 CCGCGCAGCAGGGAAGGCAAGCGCGCTCGGGAGATTTCAGCCGCTGAAGCAAG 120  
Qy 121 TGCAGATAGATGGCGTCTTAAAGATAAATCAACATTATCAAGAAAGGACAAG 180  
Db 121 TGCAGATAGATGGCGTCTTAAAGATAAATCAACATTATCAAGAAAGGACAAG 180  
Qy 181 GAACTGAAGTTGTTCAAGGAGTGTCTTGGGTCTGGTGTAGAGATCGGCTTGAATTA 240  
Db 181 GAACTGAAGTTGTTCAAGGAGTGTCTTGGGTCTGGTGTAGAGATCGGCTTGAATTA 240  
Qy 241 CCAACTGCTTTTCCCTCAGCACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
Db 241 CCAACTGCTTTTCCCTCAGCACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
Qy 301 ATCAGATGGAATATGCGGAGCCTTCCCATGTAAACATTGATCATCTTCACTGGGCT 360  
Db 301 ATCAGATGGAATATGCGGAGCCTTCCCATGTAAACATTGATCATCTTCACTGGGCT 360  
Qy 361 GGTATCAGTCACATATCTATGGCTCATTTGTTACCGGGCAGCTCTGAGCTCTCAATT 420  
Db 361 GGTATCAGTCACATATCTATGGCTCATTTGTTACCGGGCAGCTCTGAGCTCTCAATT 420  
Qy 421 GTTACCAGCATGTCATTTGAAGAATCTGTGTTCTATTTATGATCCCAATAAACTGCC 480  
Db 421 GTTACCAGCATGTCATTTGAAGAATCTGTGTTCTATTTATGATCCCAATAAACTGCC 480  
Qy 481 AAGGATCTCTCTACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAAGTTTGAAG 540  
Db 481 AAGGATCTCTCTACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAAGTTTGAAG 540  
Qy 541 AAAAGGATTTTCCCTCGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTG 600  
Db 541 AAAAGGATTTTCCCTCGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTG 600  
Qy 601 AAGAAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGTCTTAATGCGGAAC 660  
Db 601 AAGAAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGTCTTAATGCGGAAC 660  
Qy 661 AAAAGAGTCACTGTTGTCAGATAAAATGAAATTTGCTCAGCGTTGCCAGCAGCAATCA 720  
Db 661 AAAAGAGTCACTGTTGTCAGATAAAATGAAATTTGCTCAGCGTTGCCAGCAGCAATCA 720  
Qy 721 TGGGGAAGAATCTACAGTTGCTGATGACAGAGTGGATGAATGAGCCAGATATAGTTA 780  
Db 721 TGGGGAAGAATCTACAGTTGCTGATGACAGAGTGGATGAATGAGCCAGATATAGTTA 780  
Qy 781 AATACAAACACATACATGAGGAATCTAGTAAACACAGCAGCAGAAAAATCAGTATCAGC 840  
Db 781 AATACAAACACATACATGAGGAATCTAGTAAACACAGCAGCAGAAAAATCAGTATCAGC 840  
Qy 841 AGCGTCGCAGCAGGAGGAATATGACGCCAGAGCCGAGAGGAGAACCCCGCTCCCTGAGG 900  
Db 841 AGCGTCGCAGCAGGAGGAATATGACGCCAGAGCCGAGAGGAGAACCCCGCTCCCTGAGG 900  
Qy 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCCGCTGCCAGGATGACCTCGCTGCTCA 960  
Db 901 AGGACCTGTCCAAACTCTTCAAAACACACAGCCGCTGCCAGGATGACCTCGCTGCTCA 960  
Qy 961 TTGCAGGCCAGATAAAACATCTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Db 961 TTGCAGGCCAGATAAAACATCTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAACTTAG 1020  
Qy 1021 GCAAGCTCTTCATGCCCCAGGCTCTTCAAGAAATACAACTAAGAAAGGAAGTTTCCA 1080  
Db 1021 GCAAGCTCTTCATGCCCCAGGCTCTTCAAGAAATACAACTAAGAAAGGAAGTTTCCA 1080

```
QY 1081 GAAAGAGTTTAAACATGAACCTTTGAAGTCACACAGGGCAACTCTTGGAGAAATATAT 1140
Db 1081 GAAAGAGTTTAAACATGAACCTTTGAAGTCACACAGGGCAACTCTTGGAGAAATATAT 1140
QY 1141 TTGCATATTGAAAGCAGAGAGATTTCTTTAGTGTCAATGCGGATTTTGGCTATAACAG 1200
Db 1141 TTGCATATTGAAAGCAGAGAGATTTCTTTAGTGTCAATGCGGATTTTGGCTATAACAG 1200
QY 1201 TGCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260
Db 1201 TGCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260
QY 1261 AAAAAAAAAAAAAAAAAAAAA 1280
Db 1261 AAAAAAAAAAAAAAAAAAAAA 1280

RESULT 4
AX781574
LOCUS AX781574 1280 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 1 from Patent WO03050543.
ACCESSION AX781574
VERSION AX781574.1 GI:32949421
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Homo sapiens
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FEATURES
source
1.1280
/organisms="Homo sapiens"
/mol_types="unassigned DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 1280; DB 6; Length 1280;
Best Local Similarity 100.0%; Pred. No. 5e-264;
Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCGCACTCTTCCAGCTCCACCG 60
Db 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCGCACTCTTCCAGCTCCACCG 60
QY 61 CCGCGCAGCAGGGAAGCAAGGCAAGGCGGCTCGGAGATTTCAGCGGTGAAGCAAG 120
Db 61 CCGCGCAGCAGGGAAGCAAGGCAAGGCGGCTCGGAGATTTCAGCGGTGAAGCAAG 120
QY 121 TGCAGATAGATGCGCTTGTGTTAATTAAGATATCAACATTAATCAAGAGAGGCAAG 180
Db 121 TGCAGATAGATGCGCTTGTGTTAATTAAGATATCAACATTAATCAAGAGAGGCAAG 180
QY 181 GAACTGAAGTTGTTCAAGGAGTCTTTGGGCTCGGTGTGAGAGATCGGCTTGAAATTA 240
Db 181 GAACTGAAGTTGTTCAAGGAGTCTTTGGGCTCGGTGTGAGAGATCGGCTTGAAATTA 240
QY 241 CCAACTGCTTTCTTTCCCTCAGCACACAGAGATGATGCTGACITTTGATGAAGTCCAAT 300
Db 241 CCAACTGCTTTCTTTCCCTCAGCACACAGAGATGATGCTGACITTTGATGAAGTCCAAT 300
QY 301 ATCAGATGGAATATGATGCGGAGCTTTCCGCAATTAACATTAATCACTTTCACGTGGCT 360
Db 301 ATCAGATGGAATATGATGCGGAGCTTTCCGCAATTAACATTAATCACTTTCACGTGGCT 360
QY 361 GGTATCAGTCCACATATAGTATGGCTCAATCGTTACCGGGCACTCTCTGGACTCTCAGTTTA 420
Db 361 GGTATCAGTCCACATATAGTATGGCTCAATCGTTACCGGGCACTCTCTGGACTCTCAGTTTA 420
```

```
QY 421 GTTACCAGCATGCCATTGAAGAAATCTGTCTCTCTATTTATGATCCCATAAAACTGCC 480
Db 421 GTTACCAGCATGCCATTGAAGAAATCTGTCTCTCTATTTATGATCCCATAAAACTGCC 480
QY 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAAAGTTTGAAG 540
Db 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAAAGTTTGAAG 540
QY 541 AAAAGGATTTTCCCTGAAGCATTGAAAAAGCAATATCACTTTGAGTACATCTTTG 600
Db 541 AAAAGGATTTTCCCTGAAGCATTGAAAAAGCAATATCACTTTGAGTACATCTTTG 600
QY 601 AAGAAGTGGCATTTGTAATTTAAAAATTCATCTGATCAATGTCTTAATGTGGGAACCTTG 660
Db 601 AAGAAGTGGCATTTGTAATTTAAAAATTCATCTGATCAATGTCTTAATGTGGGAACCTTG 660
QY 661 AAAAGAAAGTCAGCTGTTGCAGATTAACATGAATTTGCTCAGCCTTGCCAGCAGCAATCAT 720
Db 661 AAAAGAAAGTCAGCTGTTGCAGATTAACATGAATTTGCTCAGCCTTGCCAGCAGCAATCAT 720
QY 721 TGGGGAAGAAATCTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCCAAGATATAGTTA 780
Db 721 TGGGGAAGAAATCTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCCAAGATATAGTTA 780
QY 781 AATACAAACATACATCAGGAAATCTAGTAAACAACAGCAGCAGAAACAATCAGTATCAGC 840
Db 781 AATACAAACATACATCAGGAAATCTAGTAAACAACAGCAGCAGAAACAATCAGTATCAGC 840
QY 841 AGCGTCGCCAGCAGGAGAAATATGACGCGCAGCAGCGGAGGAGAACCCCGCTCCCTGAGG 900
Db 841 AGCGTCGCCAGCAGGAGAAATATGACGCGCAGCAGCGGAGGAGAACCCCGCTCCCTGAGG 900
QY 901 AGGACTGTGCAAACTCTTCAAAACCAACACAGCCGCTGCCAGGATGGAATCGCTGCTCA 960
Db 901 AGGACTGTGCAAACTCTTCAAAACCAACACAGCCGCTGCCAGGATGGAATCGCTGCTCA 960
QY 961 TTGACAGCCAGATTAACATCACTTACTGCCAGAACATCAAGAGTTTCACTGCCCAAACTTAG 1020
Db 961 TTGACAGCCAGATTAACATCACTTACTGCCAGAACATCAAGAGTTTCACTGCCCAAACTTAG 1020
QY 1021 GCAAGCTCTTCACTGCCAGGCTCTTCAAGAAATACACAACTAAGAAAAAGGAGTTTCCA 1080
Db 1021 GCAAGCTCTTCACTGCCAGGCTCTTCAAGAAATACACAACTAAGAAAAAGGAGTTTCCA 1080
QY 1081 GAAAGAGATTAAACATGAATCTTTGAAGTCACACAGGCAACTCTTGGAGAAATATAT 1140
Db 1081 GAAAGAGATTAAACATGAATCTTTGAAGTCACACAGGCAACTCTTGGAGAAATATAT 1140
QY 1141 TTGCATATTGAAAGCAGAGAGATTTCTTTAGTGTCAATGCGGATTTTGGCTATAACAG 1200
Db 1141 TTGCATATTGAAAGCAGAGAGATTTCTTTAGTGTCAATGCGGATTTTGGCTATAACAG 1200
QY 1201 TGCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260
Db 1201 TGCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260
QY 1261 AAAAAAAAAAAAAAAAAAAAA 1280
Db 1261 AAAAAAAAAAAAAAAAAAAAA 1280
```

```
RESULT 5
HSU54559
LOCUS HSU54559 1280 bp mRNA linear PRI 08-JAN-1999
DEFINITION Homo sapiens translation initiation factor eIF3 p40 subunit mRNA,
complete cds.
ACCESSION U54559
VERSION U54559.1 GI:2351379
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE 1 (bases 1 to 1280)  
AUTHORS Asano, K., Vornlocher, H.P., Richter-Cook, N.J., Merrick, W.C.,  
Hinnebusch, A.G. and Hershey, J.W.  
TITLE Structure of cDNAs encoding human eukaryotic initiation factor 3  
subunits. Possible roles in RNA binding and macromolecular assembly  
J. Biol. Chem. 272 (43), 27042-27052 (1997)  
JOURNAL 98001678  
MEDLINE 9341143  
PUBMED  
REFERENCE 2 (bases 1 to 1280)  
AUTHORS Asano, K. and Hershey, J.W.B.  
TITLE Direct Submission  
JOURNAL Submitted (09-APR-1996) K. Asano, Biological Chemistry, University  
of California, Davis, School of Medicine, Building MSIA, Davis, CA  
95616, USA  
FEATURES  
source  
1. .1280  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue type="liver"  
6. .1064  
/note="eIF3-p40"  
/codon\_start=1  
/product="translation initiation factor eIF3 p40 subunit"  
/protein\_id="AAD03465.1"  
/db\_xref="GI:2351380"  
/translation="MAGRKETGSTATSSSTAGAAAGKGGKGGSDSAVKQVQIDGL  
VWLKIIKHVDEGGQTEVQVGLGLVVEDRLEITNCFPPQHTEDDADPEVOYQME  
MMSRLRHIDHVLGVQSYTYGSFVTRALLDSQFQYQHAIESVLLIYDPIKTAQG  
SLSLKARLPLKMEVCEKDFSEALKGAITFEYMFEEVPIVIRKSHLNLVLMWEL  
EKSVAADKHELLESLASNHLQMLMDRDEMSQDI VKYNTYMRNTSKQQQKHQ  
YQRRQENQROSRGPPLPEEDLSKLFKPPQPPARMDSLLIAGQINTYCNQIKFT  
AQNGLKFLFMAQLQAEVNN"  
ORIGIN  
Query Match 100.0%; Score 1280; DB 9; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 5e-264;  
Matches 1280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAAAGATGGCTCCCGCAAGGAAGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
DB 1 GAAAGATGGCTCCCGCAAGGAAGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
QY 61 CCGCGCAGCAGGAAAGCAAGGCAAGCGGCTCGGAGATTCAGCGCTGAAGCAAG 120  
DB 61 CCGCGCAGCAGGAAAGCAAGGCAAGCGGCTCGGAGATTCAGCGCTGAAGCAAG 120  
QY 121 TGCAGATAGATGGCTTGTGTATTAAGATAATCAACATTCAGAGAAAGCAAG 180  
DB 121 TGCAGATAGATGGCTTGTGTATTAAGATAATCAACATTCAGAGAAAGCAAG 180  
QY 181 GAACCTGAAGTTGTTCAAGGAGTCTTTGGGCTGTGGTGTAGAGAGATCGGCTTGAATTA 240  
DB 181 GAACCTGAAGTTGTTCAAGGAGTCTTTGGGCTGTGGTGTAGAGAGATCGGCTTGAATTA 240  
QY 241 CCAACTGCTTCTCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
DB 241 CCAACTGCTTCTCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
QY 301 ATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTCATTCACCTGGGCT 360  
DB 301 ATCAGATGGAATGATGCGGAGCTTCGCCATGTAACATTCATTCACCTGGGCT 360  
QY 361 GGTATCAGTCACATACCTATGCTTCATTCGTTACCCGGGCACTCTCGACTCTCAGTTA 420  
DB 361 GGTATCAGTCACATACCTATGCTTCATTCGTTACCCGGGCACTCTCGACTCTCAGTTA 420  
QY 421 GTTACCAGATGCCATGAAGAATCTGCTGTTCTCATTTATGATGCCATAAAACTGCC 480  
DB 421 GTTACCAGATGCCATGAAGAATCTGCTGTTCTCATTTATGATGCCATAAAACTGCC 480  
QY 481 AAGGATCTCTCTACTAAAGGCAATACAGATGATCTCTTAACTGATGAAGTTGTTAAAG 540  
|||||

Db 481 AAGGATCTCTCTACTAAAGGCAATACAGATGATCTCTTAACTGATGAAGTTGTTAAAG 540  
QY 541 AAAAGGATTTTCCCTCGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTG 600  
Db 541 AAAAGGATTTTCCCTCGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTG 600  
QY 601 AAGAGTGCGGATTTGTAATTTAAAAATTCAATCTGATCAATGTCTAATGTGGGAACCTTG 660  
Db 601 AAGAGTGCGGATTTGTAATTTAAAAATTCAATCTGATCAATGTCTAATGTGGGAACCTTG 660  
QY 661 AAAAGAGTCACTGCTGTCAGATAAATCAATGAATTTGCTCAGCTTGCAGCAGCAATCAT 720  
Db 661 AAAAGAGTCACTGCTGTCAGATAAATCAATGAATTTGCTCAGCTTGCAGCAGCAATCAT 720  
QY 721 TGGGGAAGATCTTACAGTTGCTGATGACAGATGGATGAATGAGCAGCAAGATATAGTTA 780  
Db 721 TGGGGAAGATCTTACAGTTGCTGATGACAGATGGATGAATGAGCAGCAAGATATAGTTA 780  
QY 781 AATACACACATACATGAGGAAATCTAGTAAACAAACAGCAGCAGCAAAATCATCAGTATCAGC 840  
Db 781 AATACACACATACATGAGGAAATCTAGTAAACAAACAGCAGCAGCAAAATCATCAGTATCAGC 840  
QY 841 AGCTGCGCCAGCAGGAATATGACGCGCCAGAGCGCGAGGAGACCCCGCTCCCTGAGG 900  
Db 841 AGCTGCGCCAGCAGGAATATGACGCGCCAGAGCGCGAGGAGACCCCGCTCCCTGAGG 900  
QY 901 AGGACCTGTCCAACTCTTCAACACACAGCAGCGCTGCCAGGATGAGCTCGTGTCTCA 960  
Db 901 AGGACCTGTCCAACTCTTCAACACACAGCAGCGCTGCCAGGATGAGCTCGTGTCTCA 960  
QY 961 TTGAGCGCAGATAAAACATTTACTGCGCAGAACATCAAGAGGATTCACCTGCCCAACTTTAG 1020  
Db 961 TTGAGCGCAGATAAAACATTTACTGCGCAGAACATCAAGAGGATTCACCTGCCCAACTTTAG 1020  
QY 1021 GCAGCTCTTCATGCGCCAGGCTCTTCAAGNATACACACATCAAGAAAGCAAGTTTCCA 1080  
Db 1021 GCAGCTCTTCATGCGCCAGGCTCTTCAAGNATACACACATCAAGAAAGCAAGTTTCCA 1080  
QY 1081 GAAAGAGGTTTAAACATGAACCTTTGAAAGTACACACAGCGGCAACTCTTGGAAAGAAATATAT 1140  
Db 1081 GAAAGAGGTTTAAACATGAACCTTTGAAAGTACACACAGCGGCAACTCTTGGAAAGAAATATAT 1140  
QY 1141 TTGCATATTTGAAAGCAGAGGATTTCTTTAGTGTCTATTCGCGATTTTGGCTATAACAG 1200  
Db 1141 TTGCATATTTGAAAGCAGAGGATTTCTTTAGTGTCTATTCGCGATTTTGGCTATAACAG 1200  
QY 1201 TGTCTTTCTAGCCATAATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260  
Db 1201 TGTCTTTCTAGCCATAATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260  
QY 1261 AAAAAAAAAAAAAAAAAAAAAA 1280  
Db 1261 AAAAAAAAAAAAAAAAAAAAAA 1280

RESULT 6  
BC000386  
LOCUS  
DEFINITION Homo sapiens eukaryotic translation initiation factor 3, subunit 3  
gamma, 40kDa, mRNA (CDNA clone MGC:8431 IMAGE:2821133), complete  
cda.  
ACCESSION BC000386  
VERSION BC000386.2 GI:38197137  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1286)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.C., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1286)

Strausberg, R.

Direct Submission

Submitted (15-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Nov. 6, 2003 this sequence version replaced gi:12653234.

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: DCTD/DP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland.

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaapi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 1 Row: e Column: 14

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503514.

# FEATURES

## source

Location/Qualifiers

1..1286

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:8431 IMAGE:2821133"

/tissue\_type="Lung, small cell carcinoma"

/clone\_lib="NIH MGC 7"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

1..1286

/gene="EIF3S3"

/notes="synonyms: eIF3-gamma, eIF3-p40"

/db\_xref="LocusID:8667"

/db\_xref="MIM:603912"

1..1059

/gene="EIF3S3"

/codon\_start=1

/product="eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa"

/protein\_id="AAH00386.1"

/db\_xref="GI:12653235"

## gene

## CDS

/db\_xref="LocusID:8667"

/db\_xref="MIM:603912"

/translation="MASRKEGTGSTATSSSTAAGKKGKSGSDSAVKQVQIDGLVLIKIKHYQEGQGVVQVGLGVVDRLEITNCFPPQHTDADDFVQYOMEVMSLRHVNIDHLHVGWYOSTYYGSFVTRALLDSQFSYQHAIEESVLLYDPIKTAGGSLSKAVRLTPKLMYCKEKDPSPEALKNTFFXMFEPVPIVKNSHLINVIMMELKKSAVADHKLHLLSSNHLGNLQLLMDRDEMDSQDIVKYNTRYMRNRSKQOQKHQYQORROENQOSRGEPLPBEEDLSKLPFPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQLQEYNN"

## ORIGIN

Query Match 98.9%; Score 1265.4; DB 9; Length 1286;  
Best Local Similarity 99.5%; Pred. No. 6.8e-261;  
Matches 1269; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 6 ATGCGCTCCGCAAGGAAGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGGC 65  
DB 1 ATGCGCTCCGCAAGGAAGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGGC 60

QY 66 GCAGCAGGGAAGGCAAGCGGCTCGGAGATTTCAGCCGTGAAGCAAGTGCAG 125  
DB 61 GCAGCAGGGAAGGCAAGCGGCTCGGAGATTTCAGCCGTGAAGCAAGTGCAG 120

QY 126 ATAGATGGCTTGTGTATTAAAGATAATCAACATTATCAAGAGAGGACAAGAACT 185  
DB 121 ATAGATGGCTTGTGTATTAAAGATAATCAACATTATCAAGAGAGGACAAGAACT 180

QY 186 GAAGTTGTTCAAGGAGTGTCTTTGGGTCTGGTTAGAGATCGGCTTGAAATATACCAAC 245  
DB 181 GAAGTTGTTCAAGGAGTGTCTTTGGGTCTGGTTAGAGATCGGCTTGAAATATACCAAC 240

QY 246 TCGTTTCTCTTCCCTCAGCACAGAGGATGATGCTGATTTGATGAAGTCAATATCAG 305  
DB 241 TCGTTTCTCTTCCCTCAGCACAGAGGATGATGCTGATTTGATGAAGTCAATATCAG 300

QY 306 ATGGAATGATCGGAGCGCTTCGCCATGTAAACATTGATCATCTTTCAGTGGCTGGTAT 365  
DB 301 ATGGAATGATCGGAGCGCTTCGCCATGTAAACATTGATCATCTTTCAGTGGCTGGTAT 360

QY 366 CAGTCCACATATATGGCTCATTCGTTACCGGGCACTCTCTGGACTCTCAGTTTAGTTAC 425  
DB 361 CAGTCCACATATATGGCTCATTCGTTACCGGGCACTCTCTGGACTCTCAGTTTAGTTAC 420

QY 426 CAGCATGCATTGAAGATCTGCTCTCATTTATGATCCATATAAACTGCCCAAGA 485  
DB 421 CAGCATGCATTGAAGATCTGCTCTCATTTATGATCCATATAAACTGCCCAAGA 480

QY 486 TCTCTCTCACTAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAG 545  
DB 481 TCTCTCTCACTAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAG 540

QY 546 GATTTTCCCTCGAAGCATTTGAAAAAGCAAAATATCACTTTGAGTACATGTTTGAAGAA 605  
DB 541 GATTTTCCCTCGAAGCATTTGAAAAAGCAAAATATCACTTTGAGTACATGTTTGAAGAA 600

QY 606 GTGCGGATTGTAATTAATAATTCATCTGATCAATGCTCTAACTGATGGAAGTTTGAAGAA 665  
DB 601 GTGCGGATTGTAATTAATAATTCATCTGATCAATGCTCTAACTGATGGAAGTTTGAAGAA 660

QY 666 AAGTCAGCTGTTGCAGATAAACAATGAAATGCTCAGCCCTTGCAGCAGCAATCATTTGGGG 725  
DB 661 AAGTCAGCTGTTGCAGATAAACAATGAAATGCTCAGCCCTTGCAGCAGCAATCATTTGGGG 720

QY 726 AAGAACTTACAGTTGCTGATGGACAGAGTGGATGAATGAGCCAAATATAGTTAAATAC 785  
DB 721 AAGAACTTACAGTTGCTGATGGACAGAGTGGATGAATGAGCCAAATATAGTTAAATAC 780

QY 786 AACACATACATGAGGAATACTAGTAAACAACAGCAGCAGAAAAACATCAGTATCAGCAGCT 845  
DB 781 AACACATACATGAGGAATACTAGTAAACAACAGCAGCAGAAAAACATCAGTATCAGCAGCT 840

QY 846 CGCCAGCAGGAGATATGACGCGCAGCGGAGGAGAACCCCGCTCCCTGAGGAGGAC 905



Db 841 CGCCAGCAGGAGAAATATCGAGCGCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGAC 900  
Qy 906 CTGTCCAAACTCTTCAAAACCAACCACAGCGCGCTGCCAGGATGGACTCGCTGCTCATTTGCA 965  
Db 901 CTGTCCAAACTCTTCAAAACCAACCACAGCGCGCTGCCAGGATGGACTCGCTGCTCATTTGCA 960  
Qy 966 GGCCAGATAAACAATTACTGCGCAGAACATCAAGGAGTTTCACTGCGCCAAAACTTtagGCAAG 1025  
Db 961 GGCCAGATAAACAATTACTGCGCAGAACATCAAGGAGTTTCACTGCGCCAAAACTTtagGCAAG 1020  
Qy 1026 CTCTTTCATGGCCCGCTCTTCAAGATACACAACTAACGAAAGGAAAGTTTCCAGAAAA 1085  
Db 1021 CTCTTTCATGGCCCGCTCTTCAAGATACACAACTAACGAAAGGAAAGTTTCCAGAAAA 1080  
Qy 1086 GAAGTTAACTGAATCTTTGAAGTCACACAGGCGCACTCTTGGAAAGAAATATATTGCA 1145  
Db 1081 GAAGTTAACTGAATCTTTGAAGTCACACAGGCGCACTCTTGGAAAGAAATATATTGCA 1140  
Qy 1146 TATTGAAAGCACAGAGATTCTTTTAGTGTCATTTGCCGATTTTGGCTATTAACAGTGCT 1205  
Db 1141 TATTGAAAGCACAGAGATTCTTTTAGTGTCATTTGCCGATTTTGGCTATTAACAGTGCT 1200  
Qy 1206 TTCTAGCCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1265  
Db 1201 TTCTAGCCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260  
Qy 1266 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1280  
Db 1261 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1275

RESULT 7  
LOCUS CO718758 1232 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 4692 from Patent WO02069579.  
ACCESSION CO718758  
VERSION CO718758.1 GI:42279615  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02069579-A 4692 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source Location/Qualifiers  
1..1232  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 96.1%; Score 1230.4; DB 6; Length 1232;  
Best Local Similarity 99.9%; Pred. No. 2.2e-253;  
Matches 1231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GAAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Db 1 GAAAGATGGCGTCCCGCAAGGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Qy 61 CCGCGCAGCAGGGAAGCAAGGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
Db 61 CCGCGCAGCAGGGAAGCAAGGCAAGCGGCTCGGAGATTCAGCCGTGAAGCAAG 120  
Qy 121 TGCAGATAGTGGCTTGTGGTATTAAAGATAATCAACATTATCAAGAGGAAGCAAG 180  
Db 121 TGCAGATAGTGGCTTGTGGTATTAAAGATAATCAACATTATCAAGAGGAAGCAAG 180  
Qy 181 GAACTGAAGTTGTTCAAGGAGTGCTTTTGGGCTGCTGGTTGAGAGATCGGCTTGAAATTA 240

Db 181 GAACTGAAGTTGTTCAAGGAGTGCTTTTGGGCTGCTGGTTAGAAAGATCGGCTTGAAATTA 240  
Qy 241 CCAACTGCTTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTTGATGAAGTCCAAT 300  
Db 241 CCAACTGCTTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTTGATGAAGTCCAAT 300  
Qy 301 ATCAGATGGAATGATGCGAGCCCTTCGCCATGTAAACATTTGATCATCTTCACGCGGCT 360  
Db 301 ATCAGATGGAATGATGCGAGCCCTTCGCCATGTAAACATTTGATCATCTTCACGCGGCT 360  
Qy 361 GGATATCAGTCCACATATATGGCTCATTCGTTACCGGGCACTCTCTGCACTCTCAGTTTA 420  
Db 361 GGATATCAGTCCACATATATGGCTCATTCGTTACCGGGCACTCTCTGCACTCTCAGTTTA 420  
Qy 421 GTTACCAGCATGCCATTTGAAGAACTCTGTCGTTCTCTATTTATGATCCCAATAAAACTGCCC 480  
Db 421 GTTACCAGCATGCCATTTGAAGAACTCTGTCGTTCTCTATTTATGATCCCAATAAAACTGCCC 480  
Qy 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAACTTTGTAAG 540  
Db 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACTGATGGAACTTTGTAAG 540  
Qy 541 AAAAGGATTTTCCCTGAAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTG 600  
Db 541 AAAAGGATTTTCCCTGAAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTG 600  
Qy 601 AAGNAGTGCCGATTTGTAATTAATAAATTCACATCTGATCAATGCTCTAATGCGAACTTG 660  
Db 601 AAGNAGTGCCGATTTGTAATTAATAAATTCACATCTGATCAATGCTCTAATGCGAACTTG 660  
Qy 661 AAAAGAGTACGCTGTTGTCAGATAAACATGAATTTGCTCAGCCTTTCAGCAGCAGCAATCA 720  
Db 661 AAAAGAGTACGCTGTTGTCAGATAAACATGAATTTGCTCAGCCTTTCAGCAGCAGCAATCA 720  
Qy 721 TGGGGAAGAACTCTACAGTCTGATGGAAGAGTGGATGAAATGAGCCAAAGATATAGTTA 780  
Db 721 TGGGGAAGAACTCTACAGTCTGATGGAAGAGTGGATGAAATGAGCCAAAGATATAGTTA 780  
Qy 781 AATAACACATACATCAGGAAATACTAGTAAACACAGCAGCAGCAAAACATCAGTATCAGC 840  
Db 781 AATAACACATACATCAGGAAATACTAGTAAACACAGCAGCAGCAAAACATCAGTATCAGC 840  
Qy 841 AGCTCGCCAGCAGGAGAAATATGAGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 900  
Db 841 AGCTCGCCAGCAGGAGAAATATGAGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 900  
Qy 901 AGGACCTGTCCAAACTCTTCAAAACCAACACAGCGCGCTGCCAGATGGAATGCTGCTCA 960  
Db 901 AGGACCTGTCCAAACTCTTCAAAACCAACACAGCGCGCTGCCAGATGGAATGCTGCTCA 960  
Qy 961 TTGCGAGCCAGATAAACAACATTTCTGCGAGAACATCAAGGAGTTTCACTGCCCCAAACTTAG 1020  
Db 961 TTGCGAGCCAGATAAACAACATTTCTACTGCGAGAACATCAAGGAGTTTCACTGCCCCAAACTTAG 1020  
Qy 1021 GCAAGCTCTTCATGGCCCGAGGCTCTTCAAGAATAACAACAATAAGAAAAAGAACTTTCCA 1080  
Db 1021 GCAAGCTCTTCATGGCCCGAGGCTCTTCAAGATAACAACAATAAGAAAAAGAACTTTCCA 1080  
Qy 1081 GAAAGAAAGTTAAACATGAACTCTTTGAAGTCAACACAGGGCAACTCTTTGGAAGAAATATAT 1140  
Db 1081 GAAAGAAAGTTAAACATGAACTCTTTGAAGTCAACACAGGGCAACTCTTTGGAAGAAATATAT 1140  
Qy 1141 TTGCTATTCAAAAGCACAGAGGATTTCTTTAGTGTCATTTGCGGATTTTGGCTATACAG 1200  
Db 1141 TTGCTATTCAAAAGCACAGAGGATTTCTTTAGTGTCATTTGCGGATTTTGGCTATACAG 1200  
Qy 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1232  
Db 1201 TGTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1232







QY 147 AAGATAATCAAAATTATCAAGAAAGCAAGAACTGAAGTTGTTCAAGAGTGCCTT 206  
DB 445 AAGATAATCAAAATTATCAAGAAAGCAAGAACTGAAGTTGTTCAAGAGTGCCTT 386  
QY 207 TTGGGTCTGGTTGTAGAAGATCGGCTTGAAATTTACCAACTGCTTTCCCTTCCCTCAGCAC 266  
DB 385 TTGGGTCTGGTTGTAGAAGATCGGCTTGAAATTTACCAACTGCTTTCCCTTCCCTCAGCAC 326  
QY 267 ACAGAGGATGATGCTGACTTTGATGAAGTCCAAATATCAGATGGAATGATCGGAGCCTT 326  
DB 325 ACAGAGGATGATGCTGACTTTGATGAAGTCCAAATATCAGATGGAATGATCGGAGCCTT 266  
QY 327 CGCCATGTAAACATTGATCATCTTCACGTGGGCTGGTATCAGTCCACATATGCTGCTCA 386  
DB 265 CGCCATGTAAACATTGATCATCTTCACGTGGGCTGGTATCAGTCCACATATGCTGCTCA 206  
QY 387 TTGCTTACCCGGGCACTCCTGACTCTCAGTTTATGTTTACCAAGCATGCCATTTGAAGAATCT 446  
DB 205 TTGCTTACCCGGGCACTCCTGACTCTCAGTTTATGTTTACCAAGCATGCCATTTGAAGAATCT 146  
QY 447 GTCGTTCTCATTTATGATCCCAATAAACTGCCCAAGATCTCTCTCACTAAAGGCATAC 506  
DB 145 GTCGTTCTCATTTATGATCCCAATAAACTGCCCAAGATCTCTCTCACTAAAGGCATAC 86  
QY 507 AGACTGACTCTTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTTGAAGCATTTG 566  
DB 85 AGACTGACTCTTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTTGAAGCATTTG 26  
QY 567 AAAAAGCAATATCACCCTTTGAGT 591  
DB 25 AAAAAGCAATATCACCCTTTGAGT 1

RESULT 14  
ID AAS25118 standard; cDNA; 783 BP.  
XX AAS25118;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human ovarian PCR-subtracted cDNA library clone #1299.  
XX  
XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;  
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;  
KW primer; probe.  
XX  
XX Homo sapiens.  
XX  
XX W0200157207-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 05-FEB-2001; 2001WO-US003733.  
XX  
XX 04-FEB-2000; 2000US-0180403P.  
XX  
XX 28-MAR-2000; 2000US-0192745P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Algate PA, Mannion J;  
XX  
XX WPI; 2001-488879/53.  
XX  
XX New polynucleotides encoding ovarian tumor proteins, useful for treating  
PT ovarian cancer, and as probes, primers, and markers of cancer  
PT progression.  
XX  
XX Example 1; Page 311-312; 378pp; English.  
XX  
XX The invention comprises compositions used for the therapy and diagnosis  
CC of ovarian cancer. The compositions comprise one or more ovarian tumour

CC proteins, their associated polynucleotides, or immunogenic portions of  
CC the proteins. The ovarian tumour polynucleotides and polypeptides are  
CC useful for stimulating and/or expanding T cells specific for a tumour  
CC protein. They are also useful for inhibiting the development of cancer in  
CC a patient with an ovarian tumour DNA or protein by incubating isolated T-  
CC cells allowing them to proliferate, and administering to the patient. The  
CC sequences can be used as markers for cancer, for example, to monitor  
CC ovarian cancer progression. Probes and primers are useful in nucleic acid  
CC hybridisation, in detecting the presence of complementary sequences in a  
CC given sample, for preparing mutant species and for preparing other  
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25238-AAS25549  
XX represent human ovarian tumour protein cDNA clones  
SQ Sequence 783 BP; 219 A; 173 C; 180 G; 196 T; 0 U; 15 Other;  
Query Match 43.9%; Score 562.4; DB 4; Length 783;  
Best Local Similarity 97.6%; Pred. No. 2.7e-94;  
Matches 603; Conservative 0; Mismatches 11; Indels 4; Gaps 3;  
QY 52 GCTCCACCCCGCGCAGCAGGAAAGGCAAGGCAAGGCGGCTCGGAGATTTCAGCCG 111  
DB 1 GCTCCACCCCGCGCAGCAGGAAAGGCAAGGCAAGGCGGCTCGGAGATTTCAGCCG 60  
QY 112 TGAAGCAAGTGCAGATAGATGCGCTTTGTGTATTAAAGATATCAACATTATCAAGAG 171  
DB 61 TGAAGCAAGTGCAGATAGATGCGCTTTGTGTATTAAAGATATCAACATTATCAAGAG 120  
QY 172 AAGGCAAGAACTGAAAGTTGTTCAAGAGTGTCTTTGGGTCTGTTGTAGAGATCGCG 231  
DB 121 AAGGCAAGAACTGAAAGTTGTTCAAGAGTGTCTTTGGGTCTGTTGTAGAGATCGCG 180  
QY 232 TTGAAATTACCAACTGCTTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 291  
DB 181 TTGAAATTACCAACTGCTTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATG 240  
QY 292 AAGTCCAATATCAGATGAAATGATGCGAGGCTTTGCGCATGTAAACATTGATCATCTTC 351  
DB 241 AAGTCCAATATCAGATGAAATGATGCGAGGCTTTGCGCATGTAAACATTGATCATCTTC 300  
QY 352 ACGTGGGTGGTATCAGTCCACATATGCTATGGCTATTTCGTTACCGGGCAGCTCTCGGACT 411  
DB 301 ACGTGGGTGGTATCAGTCCACATATGCTATGGCTATTTCGTTACCGGGCAGCTCTCGGACT 360  
QY 412 CTCAGTTTGTATACCAAGTGCCTTTGAGAACTGCTGCTCTCATTTTATGATCCCATAA 471  
DB 361 CTCAGTTTGTATACCAAGTGCCTTTGAGAACTGCTGCTCTCATTTTATGATCCCATAA 420  
QY 472 AAATGCCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAG 531  
DB 421 AAATGCCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAG 480  
QY 532 TTTGTAAAGAAAGGATTTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 591  
DB 481 TTTGTAAAGAAAGGATTTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 539  
QY 592 ACATGTTTGAAGAGAGTGCCTTGAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 648  
DB 540 ACATGTTTGAAGAGAGTGCCTTGAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 648  
QY 649 TGTGGGAACCTTGAAAAA 666  
DB 600 TGTGGGAACCTTGAAAAA 617

RESULT 15  
AAH83763  
ID AAH83763 standard; cDNA; 793 BP.  
XX AAH83763;  
XX  
XX 25-SEP-2001 (first entry)  
XX  
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:1387.

Db 1707 GACCTGTCCAAACTCTTCAAAACCAACAGCGCGCTGCCAGGATGGACTCGCTGCTCAAT 1766  
Qy 963 CGAGCCAGATAAACACTTACTGCGCAGAACATCAGGAGTTCACTGCCCCAAAACCTTAGGC 1022  
Db 1767 CGAGCCAGATAAACACTTACTGCGCAGAACATCAGGAGTTCACTGCCCCAAAACCTTAGGC 1826  
Qy 1023 AAGCTCTTTATGCGCCAGGCTCTTCAAGAAATACAACTAAGAAAAGAAAGTTTCCAGA 1082  
Db 1827 AAGCTCTTTATGCGCCAGGCTCTTCAAGAAATACAACTAAGAAAAGAAAGTTTCCAGA 1886  
Qy 1083 AAAGAAGTTACATGAACCTTTGAAGTCACACAGGCGCACTCTTGGAGAAAATATATTT 1142  
Db 1887 AAAGAAGTTACATGAACCTTTGAAGTCACACAGGCGCACTCTTGGAGAAAATATATTT 1946  
Qy 1143 GCATATTGAAAGACACAGAGATTTCTTTAGTGTCAATGCCGATTTTGGCTATACAGTG 1202  
Db 1947 GCATATTGAAAGACACAGAGATTTCTTTAGTGTCAATGCCGATTTTGGCTATACAGTG 2006  
Qy 1203 TCTTTCTAGCCATAATAAATAAAAAA 1232  
Db 2007 TCTTTCTAGCCATAATAAATAAAAAA 2036  
  
RESULT 9  
LOCUS CO413430 2285 bp DNA linear PAT 23-JAN-2004  
DEFINITION Sequence 20501 from Patent WO0170979.  
ACCESSION CO413430  
VERSION CO413430.1 GI:41321211  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
  
REFERENCE 1  
AUTHORS Lee, J. and Lillie, J.  
TITLE Genes, compositions, kits, and method for identification,  
assessment, prevention, and therapy of ovarian cancer  
JOURNAL Patent: WO 0170979-A 20501 27-SEP-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..2285  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
  
ORIGIN  
Query Match 89.0%; Score 1139.4; DB 6; Length 2285;  
Best Local Similarity 92.7%; Pred. No. 7.4e-234;  
Matches 1248; Conservative 0; Mismatches 26; Indels 73; Gaps 2;  
  
Qy 1 GAAAGATGGCTCC CGCAAGGAAGGTACCGGCTCTACTGCACTCTTCCAGCTCCACC 59  
Db 37 GAAAGATGGCTCCGCGCAGGAAGGTACCGGCTCTACTGCACTCTTCCAGCTCCACC 96  
Qy 60 GCCGCGCAGCAGGAAGGCAAGGCAAGCGCGCTCGGAGATTCAAGCGTCAAGCAA 119  
Db 97 GCCGCGCAGCAGGAAGGCAAGGCAAGCGCGCTCGGAGATTCAAGCGTGAAGCAA 156  
Qy 120 GTGCAGATAGATGGCTTGTGGTATTAAGATAATCAACATTAATCAAGAGATTCAGGAGAGAGCAA 179  
Db 157 GTGCAGATAGATGGCTTGTGGTATTAAGATAATCAACATTAATCAAGAGATTCAGGAGAGAGCAA 216  
Qy 180 GGAAGTGAAGTTGTTCAAGAGATGCTTTTGGTCTGGTTGTAGAGATTCGGCTTGAATTT 239  
Db 217 GGAAGTGAAGTTGTTCAAGAGATGCTTTTGGTCTGGTTGTAGAGATTCGGCTTGAATTT 276  
Qy 240 ACCAAGTCTTTCTTTCCCTCAGCACACAGAGATGATGCTGACTTTTGATGAGAG----- 294  
Db 277 ACCAAGTCTTTCTTTCCCTCAGCACACAGAGATGATGCTGACTTTTGATGAGAGTTCA 336  
Qy 295 ----- 294

Db 337 TTCATAGTCAGCATACATCAGAACTTTCGTTCTCTGTTATTGCTGAGTAGTATTCCATTA 396  
Qy 295 -----TCCAATATCAGATCGAAATCATGTCGGAGGCTTCGCCATGTAAACATTTGATCAT 347  
Db 397 AATGTTATTCATATCAGATCGAAATCATGTCGGAGGCTTCGCCATGTAAACATTTGATCAT 456  
Qy 348 CTTCCAGCTGGGCTGGTATCAGTCCACATATATATGGCTCATTTGTTACCGGGGCACTCCTG 407  
Db 457 CTTCCAGCTGGGCTGGTATCAGTCCACATATATATGGCTCATTTGTTACCGGGGCACTCCTG 516  
Qy 408 GACTCTCAGTTTAGTTACAGCATGCCATTTGAAGAATCTGTCGTTCTCATTTATGATCCC 467  
Db 517 GACTCTCAGTTTAGTTACAGCATGCCATTTGAAGAATCTGTCGTTCTCATTTATGATCCC 576  
Qy 468 ATAAAACTGCCCAAGGATCTCTCTCACTAAGGCATACAGACTGACTCCCTAAACTGATG 527  
Db 577 ATAAAACTGCCCAAGGATCTCTCTCACTAAGGCATACAGACTGACTCCCTAAACTGATG 636  
Qy 528 GAAGTTTGTAAAGAAAAGGATTTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTT 587  
Db 637 GAAGTTTGTAAAGAAAAGGATTTTTCCCTGAAGCATTTGAAAAAGCAAAATATCACCTTT 696  
Qy 588 GAGTACATGTTTGAAGAGTGGCCGATTTGTAATTTAAAAATTCACATCTGATCAATGTCTTA 647  
Db 697 GAGTACATGTTTGAAGAGTGGCCGATTTGTAATTTAAAAATTCACATCTGATCAATGTCTTA 756  
Qy 648 ATGTGGCACTTGAAGAGAGTGCAGTGTTCAGATAAACATGAATTTGCTCAGCTTCGCC 707  
Db 757 ATGTGGCACTTGAAGAGAGTGCAGTGTTCAGATAAACATGAATTTGCTCAGCTTCGCC 816  
Qy 708 AGCAGCAATCATTTGGGGAAGAACTTACAGTTGCTGATGAGCAGAGTGGATGAATGAGC 767  
Db 817 AGCAGCAATCATTTGGGGAAGAACTTACAGTTGCTGATGAGCAGAGTGGATGAATGAGC 876  
Qy 768 CAAGATATAGTTAAATACACATACATGAGGAATCTAGTAAACACAGCAGCAGAGAAA 827  
Db 877 CAAGATATAGTTAAATACACATACATGAGGAATCTAGTAAACACACAGCAGCAGAGAAA 936  
Qy 828 CATCAGTATCAGCAGCTCGCCAGCAGAGAAATATGAGCGCCAGAGCCGAGGAGAGAACCC 887  
Db 937 CATCAGTATCAGCAGCTCGCCAGCAGAGAAATATGAGCGCCAGAGCCGAGGAGAGAACCC 996  
Qy 888 CCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCCAAACCAACACAGCGCGCTGCCAGGATG 947  
Db 997 CCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCCAAACCAACACAGCGCGCTGCCAGGATG 1056  
Qy 948 GACTCGCTGCTCATTTGAGGCGCAGATAAACATTTACTGCGCAGAAATCAACAGGAGTTCACT 1007  
Db 1057 GACTCGCTGCTCATTTGAGGCGCAGATAAACATTTACTGCGCAGAAATCAACAGGAGTTCACT 1116  
Qy 1008 GCCCAAACTTTAGCAGCTCTTTCATGCGCCGCGCTCTTCAAGATACCAACATTAAGAA 1067  
Db 1117 GCCCAAACTTTAGCAGCTCTTTCATGCGCCGCGCTCTTCAAGATACCAACATTAAGAA 1176  
Qy 1068 AAGGAAGTTTCCAGAAAAGGATTAACATGAATCTTTGAAGTCAACACAGGCGCAACTCTT 1127  
Db 1177 AAGGAAGTTTCCAGAAAAGGATTAACATGAATCTTTGAAGTCAACACAGGCGCAACTCTT 1236  
Qy 1128 GGAAGAAATATATTTGATATTTGAAAAAGCAGAGGATTTCTTTAGTGTCAATGCCGAT 1187  
Db 1237 GGAAGAAATATATTTGATATTTGAAAAAGCAGAGGATTTCTTTAGTGTCAATGCCGAT 1296  
Qy 1188 TTGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 1247  
Db 1297 TTGCTATAACAGTGTCTTTCTAGCCATAATAAATAAAAAA 1356  
Qy 1248 AAAAAAAAAAAAAAAAAAAAAA 1274  
Db 1357 CAGCAACACACAGCACCACCAAAA 1383  
  
RESULT 10

CR456906	1059 bp	mRNA	linear	PRI 03-JUN-2004
LOCUS				
DEFINITION	Homo sapiens full open reading frame cDNA clone RZPD0834F0211D for gene EIF3S3, eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa; complete cds, incl. stopcodon.			
ACCESSION	CR456906.1	GI:48145928		
VERSION	Full ORF shuttle clone, Gateway(TM), complete cds.			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1059)			
AUTHORS	Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.			
TITLE	Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1059)			
AUTHORS	Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.			
TITLE	Direct Submission			
JOURNAL	Submitted (03-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany			
COMMENT	RZPD: RZPD0834F0211D, ORFNo 1254 www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPD0834F0211D RZPDLIB; Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No. 834 www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834 www.rzpd.de/products/orfclones/ Contact: Inge Arlart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 100 Fax: +49 30 32639 111 www.rzpd.de This clone is available from RZPD; contact RZPD (customer.service@rzpd.de) for further information. This CDS clone is a part of a collection of human full length expression clones generated by RZPD. This CDS has been cloned incl. stopcodon. The CDS has been inserted into pDONR201 via a BP Clonase(TM) reaction. Additional sequence has been added in front of the start codon: att: .AAAAA GCA GGC (ATG). The last base of the last coding triplet has been changed to T, which might lead to an amino acid change at the C terminus of the polypeptide. The stop codon has been set to TAA followed by TTAACCCAGTTCTT. att. Compared to the reference sequence NM_003756 we did not find any amino acid exchanges. Clone distribution: http://www.rzpd.de/products/orfclones/. Location/Qualifiers 1. 1059 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="RZPD0834F0211D" /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD" /lab_host="DH10B" /note="Vector: pDONR201, Site_1: attp1; Site_2: attp2" 1. 1059 /genes="EIF3S3" 1. 1059 /genes="EIF3S3" /codon_start=1 /protein_id="CAG33187.1" /db_xref="GI:48145929" /translations="MASRKEGTGSTATSSSTAGAGKKGKGGSDNSAKVQVQIDGLVILKIKHVEECQTEVQVQVLLGLVVEDRLREINCFPPORTEDDADFDEVQOME MRSLSKHVDILHWGVSQTYGSFVTRALLDSQFSYOHAIIEESVWLYLIDFKTAQG SLSKAYRITPLKMEVCKKQFSPEALKKANTTFEYFEEVPIVKNLHNLVLMWEL EKKSAYADHELKSLASSNHLGNLQMLDRPDEMDSQDITVKNYTRNTSKQOQKHQ YQORQOENQRQSRGPEPLPBEDLSKLFKPPPPARMDSLIAQIINTYCQNIKEFT AQNLGKFLNAQALQYNN" 725 720 785 780 845 840 905 900 965 960 1025 1020			

## ORIGIN

	Query Match	82.6%	Score 1057.4;	DB 9;	Length 1059;
	Best Local Similarity	99.9%	Pred. No. 2.7e-216;		
	Matches 1058;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
QY	6	ATGCGCTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACGCGGCG	65		
Db	1	ATGCGCTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACGCGGCG	60		
QY	66	GCAGCAGGAAAGCAAGCGCGCTCGGAGATTTCAGCCCTGAAGCAAGTGCAG	125		
Db	61	GCAGCAGGAAAGCAAGCGCGCTCGGAGATTTCAGCCCTGAAGCAAGTGCAG	120		
QY	126	ATGATGCGCTTGTGTATTAAAGATAATCAACATTATCAAGAAAGCAAGAACT	185		
Db	121	ATGATGCGCTTGTGTATTAAAGATAATCAACATTATCAAGAAAGCAAGAACT	180		
QY	186	GAAGTTGTTCAAGGAGTCTTTTGGGTCTGGTTGTAGAAGATCGGCTTCAAAATTACCAAC	245		
Db	181	GAAGTTGTTCAAGGAGTCTTTTGGGTCTGGTTGTAGAAGATCGGCTTCAAAATTACCAAC	240		
QY	246	TGCTTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTTGATGAAGTCCAATATCAG	305		
Db	241	TGCTTTCCCTTCCCTCAGCACACAGAGGATGATGCTGACTTTTGATGAAGTCCAATATCAG	300		
QY	306	ATGGAATGATCGGAGCGCTTCGCCATGTAAACATTGATCATCTTCACTGGCTGGTAT	365		
Db	301	ATGGAATGATCGGAGCGCTTCGCCATGTAAACATTGATCATCTTCACTGGCTGGTAT	360		
QY	366	CAGTCACACATCTATGGCTCATTCCTCCGGGACCTCTCGACTCTCAGTTTATGTTAC	425		
Db	361	CAGTCACACATCTATGGCTCATTCCTCCGGGACCTCTCGACTCTCAGTTTATGTTAC	420		
QY	426	CAGCATGCCATTGAAGAATCTGCTCTCTCAATTTATGATCCATAAAAACTGCCCAAGA	485		
Db	421	CAGCATGCCATTGAAGAATCTGCTCTCTCAATTTATGATCCATAAAAACTGCCCAAGA	480		
QY	486	TCCTCTCATAAAGGCATACAGCTGACTCTTAACCTGATGGAAATTTGTAAGAAAG	545		
Db	481	TCTCTCTACTAAGGCATACAGCTGACTCTTAACCTGATGGAAATTTGTAAGAAAG	540		
QY	546	GATTTTCCCTTGAAAGCATTCGAAAAAGCAAAATATCACTTTGAGTACATGTTTCAAGA	605		
Db	541	GATTTTCCCTTGAAAGCATTCGAAAAAGCAAAATATCACTTTGAGTACATGTTTCAAGA	600		
QY	606	GTGCGGATGTAATTAATAATTCATCTGATCAATGCTTAAATGTGGGAACCTGAAAG	665		
Db	601	GTGCGGATGTAATTAATAATTCATCTGATCAATGCTTAAATGTGGGAACCTGAAAG	660		
QY	666	AAATGAGCTGTTGCAGATAAAACATGAATTTGCTCAGCCCTTGCAGCAGCAATCATTTGGG	725		
Db	661	AAATGAGCTGTTGCAGATAAAACATGAATTTGCTCAGCCCTTGCAGCAGCAATCATTTGGG	720		
QY	726	AAGATCTACAGTTGCTGATGGACAGAGTGGATGAATGAGCAAGATATAGTTAAATAC	785		
Db	721	AAGATCTACAGTTGCTGATGGACAGAGTGGATGAATGAGCAAGATATAGTTAAATAC	780		
QY	786	AACACATACATGAGGAATCTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGGT	845		
Db	781	AACACATACATGAGGAATCTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGGT	840		
QY	846	CGCCAGCAGGAATATGCGCGCCAGAGCGGAGAGAACCCCGCTCCCTCAGGAGGAC	905		
Db	841	CGCCAGCAGGAATATGCGCGCCAGAGCGGAGAGAACCCCGCTCCCTCAGGAGGAC	900		
QY	906	CTGTCCAAACTCTTCAAACACCAACACAGCGGCTGCCAGGATGGACTCGTCTCATTTGCA	965		
Db	901	CTGTCCAAACTCTTCAAACACCAACACAGCGGCTGCCAGGATGGACTCGTCTCATTTGCA	960		
QY	966	GGCAGATAAACACTTACTGCGCAACATCAAGAGTTCCTGCGCCCAAACTTAGGCAAG	1025		
Db	961	GGCAGATAAACACTTACTGCGCAACATCAAGAGTTCCTGCGCCCAAACTTAGGCAAG	1020		

FEATURES  
sourcegene  
CDS

Qy 1026 CTCTTTCAGCCAGGCTCTTCAAGAAATACAACTAA 1064  
Db 1021 CTCTTTCAGCCAGGCTCTTCAAGAAATACAACTAA 1059

RESULT 11  
AX774278  
LOCUS AX774278 1045 bp DNA linear PAT 09-JUL-2003  
DEFINITION Sequence 22 from Patent W003046176.  
ACCESSION AX774278  
VERSION AX774278.1 GI:32485957  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Legrain, P., Rain, J.C., Benarous, R., Emiliani, S., Berlioz-Torrent, C.  
and Blot, G.  
TITLE Protein-protein interactions in human immunodeficiency virus  
JOURNAL Patent: W0 03046176-A 22 05-JUN-2003;  
Hybrigenics (FR)  
FEATURES  
source location/Qualifiers  
1..1045  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 81.6%; Score 1045; DB 6; Length 1045;  
Best Local Similarity 100.0%; Pred. No. 1.2e-213;  
Matches 1045; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 GGAAGGTACCGCTCTACTGACCTCTTCAGCTCCACCGCGGAGGAAAGG 79  
Db 1 GGAAGGTACCGCTCTACTGACCTCTTCAGCTCCACCGCGGAGGAAAGG 60

Qy 80 CAAAGGCAAGCGGCTCGGAGATTGAGGAGTTCAGCGTGAAGTGCAGATAGATGCTTGT 139  
Db 61 CAAAGGCAAGCGGCTCGGAGATTGAGGAGTTCAGCGTGAAGTGCAGATAGATGCTTGT 120

Qy 140 GGTTAATGAATATCAACATTTATCAAGAGAGGACAGGAATCAAGTGTTCAGG 199  
Db 121 GGTTAATGAATATCAACATTTATCAAGAGAGGACAGGAATCAAGTGTTCAGG 180

Qy 200 AGTCTTTGGTCTGGTTGAGAGATCGCTTGAATACCACTGCTTCCTTCCC 259  
Db 181 AGTCTTTGGTCTGGTTGAGAGATCGCTTGAATACCACTGCTTCCTTCCC 240

Qy 260 TCAGCACAGAGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATGATGCG 319  
Db 241 TCAGCACAGAGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATGATGCG 300

Qy 320 GAGCTTCGCGCATGTAACATTTGATCATCTTCAGTGGGCTGGTATCAGTCCACATACTA 379  
Db 301 GAGCTTCGCGCATGTAACATTTGATCATCTTCAGTGGGCTGGTATCAGTCCACATACTA 360

Qy 380 TGGCTCATTTGTTACCGGGACCTCTGGAATCTCAGTTAGTTACAGCATGCCATGTA 439  
Db 361 TGGCTCATTTGTTACCGGGACCTCTGGAATCTCAGTTAGTTACAGCATGCCATGTA 420

Qy 440 AGAATCTGCTCTCTCATTTATGATCCATATAAAGTCCCAAGGATCTCTCACTAAA 499  
Db 421 AGAATCTGCTCTCTCATTTATGATCCATATAAAGTCCCAAGGATCTCTCACTAAA 480

Qy 500 GGCTATACAGTGAATCTCTAACTGATGGAAGTTTGTAAAGAAAAGGATTTTCCCTGA 559  
Db 481 GGCTATACAGTGAATCTCTAACTGATGGAAGTTTGTAAAGAAAAGGATTTTCCCTGA 540

Qy 560 AGCATTGAAGAAAAGCAATATCACTTTTGATGATCATGTTGAAGAAAGTGCAGATTGAAT 619  
Db 541 AGCATTGAAGAAAAGCAATATCACTTTTGATGATCATGTTGAAGAAAGTGCAGATTGAAT 600

Qy 620 TAAAAATTCATCTGATCAATGTCTTAATGTGGAACTTGGAAAGAGTCAAGCTGTTC 679  
Db 601 TAAAAATTCATCTGATCAATGTCTTAATGTGGAACTTGGAAAGAGTCAAGCTGTTC 660

Qy 680 AGATAAATCATGAAATTCAGCTTCAGCTTCAGCAGAGCAATCATTTGGGGAAGATCTACAGTT 739  
Db 661 AGATAAATCATGAAATTCAGCTTCAGCTTCAGCAGAGCAATCATTTGGGGAAGATCTACAGTT 720

Qy 740 GCTGATCGACAGAGTGTGAAATGAGCAAGATATAGTTAAATACAAACATACATGAG 799  
Db 721 GCTGATCGACAGAGTGTGAAATGAGCAAGATATAGTTAAATACAAACATACATGAG 780

Qy 800 GAATACTAGTAAACAAACAGCAGCAGCAAAACATCAGTATCAGCAGCGTCCCGCAGCAGGAGAA 859  
Db 781 GAATACTAGTAAACAAACAGCAGCAGCAAAACATCAGTATCAGCAGCGTCCCGCAGCAGGAGAA 840

Qy 860 TATGACGGCCAGAGCGGAGAGAACCCCGCTCCCTGAGGAGGACCTGTCCAACTCTT 919  
Db 841 TATGACGGCCAGAGCGGAGAGAACCCCGCTCCCTGAGGAGGACCTGTCCAACTCTT 900

Qy 920 CAAACACACAGCGCGCTGCGAGGATGAGTCTGCTCATTCAGCAGCCAGATAAACAC 979  
Db 901 CAAACACACAGCGCGCTGCGAGGATGAGTCTGCTCATTCAGCAGCCAGATAAACAC 960

Qy 980 TTACTGCCAGAACATCAAGGAGTTCACTGCGCCAAAACCTTAGGCAAGCTCTTTCATGGCCCA 1039  
Db 961 TTACTGCCAGAACATCAAGGAGTTCACTGCGCCAAAACCTTAGGCAAGCTCTTTCATGGCCCA 1020

Qy 1040 GGCTCTTCAAGATATCAACAACTAA 1064  
Db 1021 GGCTCTTCAAGATATCAACAACTAA 1045

RESULT 12  
BC060586 1273 bp mRNA linear ROD 30-JUN-2004  
LOCUS Rattus norvegicus eukaryotic translation initiation factor 3,  
DEFINITION subunit 3 gamma, 40kDa, mRNA (cdna clone MGC:72941 IMAGE:6918037),  
complete cds.  
BC060586  
ACCESSION BC060586.1 GI:38051897  
VERSION Rattus norvegicus (Norway rat)  
KEYWORDS Rattus norvegicus  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1273)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Shapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywicki, M.I., Skalski, U., Smailus, D.E.,  
Schnarch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
JOURNAL 12477932  
PUBMED 12477932  
REFERENCE 2 (bases 1 to 1273)  
AUTHORS Strausberg, R.

**TITLE**  
**JOURNAL**  
 Submitted (27-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**REMARK**  
**COMMENT**  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: John C. Marshall, M.D., Ph.D  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska, Duane Smalus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 52 Row: 9 Column: 23  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, Similarity but not identity to protein.

#### FEATURES

source

Location/Qualifiers  
 1..1273  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="MGC:72941 IMAGE:6918037"  
 /tissue\_type="Pituitary gland, anterior, rat"  
 /clone\_lib="NICHDR\_Pit1"  
 /lab\_host="DH10B"  
 /notes="vector: pDNR-LIB"  
 1..1273  
 /gene="Elf3s3"  
 /notes="synonym: MGC72941"  
 /db\_xref="LocusID:299899"  
 /db\_xref="RGD:735178"  
 30..1088  
 /gene="Elf3s3"  
 /codon\_start=1  
 /product="eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa"  
 /protein\_id="AAH60586.1"  
 /db\_xref="GI:38051898"  
 /db\_xref="LocusID:299899"  
 /db\_xref="RGD:735178"

gene

CDS

#### ORIGIN

Query Match 73.1%; Score 935.8; DB 10; Length 1273;  
 Best Local Similarity 85.7%; Pred. No. 3.3e-190;  
 Matches 1077; Conservative 0; Mismatches 172; Indels 8; Gaps 3;

QY 1 GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCCACTCTTCAGCTCCACCG 60  
 DB 25 GAAACATGGCGTCGCGCAAGGAGGACCGGTTCCACGCCACCTCTTCAGCTCTACTG 84  
 QY 61 CCGGCCAGCAGGAAGCAAGGCAAGCGCGCTCGGAGATTGAGCGGTGAAGCAAG 120

DB 85 GCGGCGGCTGGGAAGGAAAAGGCGGCTCTCGAGATTCGGCGCTGAAGCAGG 144  
 QY 121 TGCAGATAGATGGCTTGTGGTATTAAAGATTAATCAAACTATCAAGAAAGAGCAAG 180  
 DB 145 TGCAGATTGACGGCTGTGGTATTAAAGATTAATCAAACTATCAAGAAAGAGCAAG 204  
 QY 181 GAACTGAAGTGTTCGAAGGAGTCTTTGGGTCTGGTCTGTAGAAAGATCGCTTGAATTA 240  
 DB 205 GAACTGAAGTGTTCGAAGGAGTCTCTGGGCTGTGGTGTAGAAAGACCGAGATTA 264  
 QY 241 CCAACTGCTTTCTTTCCCTTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAAT 300  
 DB 265 CCAACTGTTTCCCACTTTCCCCAGCACACAGAGGAGCGATGCTGACTTTGATGAAGTCA 324  
 QY 301 ATCAGATGAAGATGATCGGAGCGCTTGGCCATGTAAACATTTGATCATCTTCAGTGGGCT 360  
 DB 325 ATCAGATGAGATGATGCGCAGCGCTTGGCCAGCGTGAACATCGACCACTCCACGTTGG 384  
 QY 361 GGTATCAGTCCACATATGCTCATTTGTTACCGGGCACCTCTCTGGACTCTCTCAGTTTA 420  
 DB 385 GGTACAGTCCAGATTAATGCTCTCTGTTACCGGGCGCTCTTGGATTTCTCAGTTCA 444  
 QY 421 GTTACAGCATGCCATTGAAGAAATCTGCTGTTCTCATTTATGATCCCAATAAACTGCC 480  
 DB 445 GCTACAGCAGCGCACTTGAAGAAATCTGCTGTTCTCATTTATGATCCCAATAAACTGCC 504  
 QY 481 AAGGATCTCTCACTAAAGGATACAGACTGACTCTTAACTGATGGAAGTTTGTAAAG 540  
 DB 505 AAGGATCTCTCACTAAAGGATACAGACTGACTCTTAACTGATGGAAGTTTGTAAAG 564  
 QY 541 AAAAGGATTTTCCCTTGAAAGCAATGMAAAAGCAAAATACCTTTGAGTACATGTTTG 600  
 DB 565 AGAAGGACTTTTCCCGGAGGATTTGMAAAAGCAAAATACCTTTGAGTACATGTTTG 624  
 QY 601 AAGAAAGTCGATTTGTAATTAATAATTCATCTGATCAATCTCTTAATGTGGGAATTTG 660  
 DB 625 AAGAAAGTCGATTTGTAATTAATAATTCATCTGATCAATCTCTTAATGTGGGAATTTG 684  
 QY 661 AAAAGAGTCAGCTGTTCGAGATAACATGATTTGCTGAGCTTCCAGCAGCAATCAT 720  
 DB 685 AGAAGAGTCAGCTGTTCGAGATAACATGATTTGCTGAGCTTCCAGCAGCAATCAT 744  
 QY 721 TGGGAAGAAATCTACAGTTGCTGATGGAGCAGAGTGGATGAAATGAGCCAAAGATATAG 780  
 DB 745 TGGGAAGAAATCTGAGTTGCTGATGGAGCAGAGTGGATGAAATGAGTCAAGACATA 804  
 QY 781 AATACAACATACATAGAGGAATACTAGTAAACAAACAGCAGCAGAAACATCAGTATCAG 840  
 DB 805 AATACAACATACATAGAGGAATACTAGTAAACAAACAGCAGCAGAAACATCAGTATCAG 864  
 QY 841 AGCGTCGCCAGCAGAGATATGCGCGCCAGAGCGGAGGAGAAACCCCGCTCCCTGAGG 900  
 DB 865 AGCGTCGCCAGCAGAGATATGCGCGCCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 924  
 QY 901 AGGAAGCTGCAAACTCTTCAAAACCAACACAGCCGCTGCCAGGATGAGTCTCGCTCA 960  
 DB 925 AGGAAGCTGCAAACTCTTCAAGCCCAACAGCCGCTGCCAGGATGAGTCTCGCTCA 984  
 QY 961 TTGAGGCGCAGATAAACCTTACTGCCAGAACATCAAGAGGTTCTACTGCCAAAACTTAG 1020  
 DB 985 TTGAGGCGCAGATAAACCTTACTGCCAGAACATCAAGAGGTTCTACTGCCAAAACTTAG 1044  
 QY 1021 GCAAGCTCTTATGCCCGCAGGCTCTTCAAGATCAACAACTAAGAAAGGAGTTTCCA 1080  
 DB 1045 GCAAACTCTTATGCTCAGGCTCTTCAAGATCAAGTAACTAAGAAAGGAGGAGTCTCT 1104  
 QY 1081 GAAAAAGAGTTAACTGAATCTTGAAGTCAACAGGCGCACTCTTGAAGAAATATAT 1140  
 DB 1105 GCAAGAGG-----ACAGACTGTTGAGACTCCACAGAGTGAATCTTGA--AAATGTGT 1157  
 QY 1141 TTGCATATTGAAAGCAGACAGGATTTCTTTAGTGTCTATGCCGATTTTGGCTATAACAG 1200  
 DB 1158 TTCTGTACTGCAAGTGGAGGCGCACTGTGCTCTTAGTGTG- TAATGGCTTTATAAC 1216



Qy 1201 TGTCTTTCTAGCCATAATAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1257  
 Db 1217 CAATAATAAAGAGTTACTTGGTGTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1273

RESULT 13  
 BC014755 1254 bp mRNA linear ROD 29-JUN-2004  
 LOCUS Mus musculus eukaryotic translation initiation factor 3, subunit 3  
 (gamma), mRNA (cdna clone MGC:25493 IMAGE:4504013), complete cds.  
 ACCESSION BC014755  
 VERSION BC014755.1 GI:15928557  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1254)  
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,  
 Villaalon, K.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalusz, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932  
 2 (bases 1 to 1254)  
 Strausberg, R.  
 Direct Submission  
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NTH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: The Corio Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 31 Row: 1 Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 18079340.

FEATURES  
 Location/Qualifiers  
 1..1254  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"

gene  
 1..1254  
 /gene="Eif3a3"  
 /note="synonyms: EIF3-P40, EIF3-gamma, 40kD"  
 /db\_xref="LocusID:68135"  
 /db\_xref="MGI:1915385"  
 5..1063  
 /gene="Eif3a3"  
 /codon\_start=1  
 /product="eukaryotic translation initiation factor 3,  
 subunit 3 (gamma)"  
 /protein\_id="AAH14755.1"  
 /db\_xref="GI:15928558"  
 /db\_xref="LocusID:68135"  
 /db\_xref="MGI:1915385"  
 /translation="MASRKEGTSTATSSGAGGAVGKGGSGDSAVKQVQIDGL  
 MMRSLRVNIDHLHVGWQSYTYGSFVTRALLDSQFSYQHAIBESVLLIYDPIKTAQG  
 SLSLKAYRLTPKLMVECKEDPSPALKASITEHPPEVPIVKKSHLNLVLMWEL  
 EKSSAVADKHELUSLASSNHLGKLMKRVDEMSSQDIIKINITYMNTSKQQQKHQ  
 YQRRQENMQRSGRPPLPEEDLSKLFKPHQAPARMDSLLIAGQINTYQNTKEFT  
 AQNLGKLFMAQALQAEYNN"

ORIGIN  
 Query Match 71.1%; Score 910; DB 10; Length 1254;  
 Best Local Similarity 82.9%; Pred. No. 1.le-184;  
 Matches 1039; Conservative 0; Mismatches 215; Indels 0; Gaps 0;  
 Qy 2 AAACATGCGCTCCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGTCCACGCG 61  
 Db 1 AAACATGCGCTCCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGTCCACGCG 60  
 Qy 62 CGGCGCAGCAGGAAAGGCAAGGCAAGGCGGCTCGGAGATTGAGCGGTGAAGCAAGT 121  
 Db 61 CGGCGCGTGGGAAAGGCAAGGCAAGGCGGCTCGGAGATTGAGCGGTGAAGCAAGT 120  
 Qy 122 GCAGATAGATGGCTTGTGTATTAAGATTAATCAACATTTATCAAGNAGAGGACAGG 181  
 Db 121 GCAGATCGACGGCTGTGTATTAAGATTAATCAACATTTATCAAGNAGAGGACAGG 180  
 Qy 182 AACTGAATGTTTCAAGGAGTGCTTTGGGTCTGTTGTAGAAGATCGGCTTGAATTTAC 241  
 Db 181 CACTGAGTTTTCAGGGCGTGCTCTCGGTCTGTTGTGGAAGACCGGCTAGAGATTAC 240  
 Qy 242 CAACGTCTTCTTCCCTCAGCACACAGAGGATGCTGCTGACTTTGATGAAGTCCAATA 301  
 Db 241 CAACGTCTTCCCGTTCCCGCAGCACAGGAGGATGATGCTGACTTTGATGAGGTACAGTA 300  
 Qy 302 TCAGATGGAATGATGCGGAGCCTTCGCCATGTAACATTTGATCATCTTACGTGGGCTG 361  
 Db 301 TCAGATGGAATGATGCGGAGCCTTCGCCATGTAACATTTGATCATCTTACGTGGGCTG 360  
 Qy 362 GTATCAGTCCACATCTATGCTCATTTGTTACCGGCGCTCTCTGACTCTCAGTTTAG 421  
 Db 361 GTATCAGTCCACATTTACCGCTCTCTGTTTACCGGCGGCTCTGAGATTTCTAGTTTCA 420  
 Qy 422 TTACAGCATGCCATTGAAGAATCTGCTGTTCTCATTTATGATCCCAATAAACTGCCCA 481  
 Db 421 CTACAGCAGCCATTGAAGATCTGCTGCTCTCATTTATGATCCCAATAAACTGCCCA 480  
 Qy 482 AGGATCTCTCACTAAGGCATACAGACTGATCTCTAACTGATGAAGTTTGAAGA 541  
 Db 481 AGGATCTCTCGCTGAAGCGGTACAGACTGACTCTCTAACTGATGAAGTTTGAAGA 540  
 Qy 542 AAAGGATTTTCCCTCAAGCATTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGA 601  
 Db 541 GAAGGACTTTTCCCTCAAGCATTGAAAAAGCAAGCATCACCTTTGAGCACAATGTTGA 600  
 Qy 602 AGAGTCCGATTGTGTAATTAATAATTCACATCTGATCAATGTCCTTAATGTGGGAACTTGA 661



Db 601 AGAAGTCCGATTTGTAATTAATAAACTCACATCTGATCAATCTCTTATGTGGGAATTGA 660  
Qy 662 AAAGAGTCAGCTGTTGCAGATAAACAATGAATTCGTCTGACCTTCCGACGAGCAATCAATTT 721  
Db 661 GAAGAGTCAGCTGTTGGCGGTAAGACACGAATTCGTCTGAGTCTTCTAGCAGCAATCATCT 720  
Qy 722 GGGGAAGAATCTACAGTTGCTGTGAGCAGAGTGGATGAATGAGCCCAAGATATAGTTAA 781  
Db 721 GGGGAAGAGCTTCAGCTGCTGATGGACCGGGTGGACGAATGAGCCAGGACATATCAAA 780  
Qy 782 ATACAACACATACATGAGGAATCTAGTAAACAACAGCAGCAGAGCAATCAGTATCAGCA 841  
Db 781 ATACAACACGTACATGCGCAACACAGTAAGCAGCAGCAGCAGAAAATCAGTATCAGCA 840  
Qy 842 GCGTCGCAGCAGAGAGAAATATGACGCGCAGAGCGGAGGAGAACCCCGCTCCCTGAGGA 901  
Db 841 GCGTCGCACACAGGAGAAATATGACGCGCAGAGTCAGAGTCGAGCGAGCCGCCACTCCCTGAGGA 900  
Qy 902 GGACTGTCCAAACTCTTCAAACACACACAGCGCCCTGCCAGGATGACCTGCTGCTCAT 961  
Db 901 GGAACCTCTTAAACTCTTCAAGCCCACAGGCCCTGCCAGGATGGAATCACTGCTCAT 960  
Qy 962 TGCAGGCAGATAAACAATCTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAACTTAGG 1021  
Db 961 TGCAGGCAGATAAACAATCTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAACTTAGG 1020  
Qy 1022 CAGCTCTTCATGCGCCAGGCTCTTCAAGATACACAACTAGAAAGAGAGTTTCCAG 1081  
Db 1021 CAAACTCTTATGCGCCAGGCTCTTCAAGATACAACTAGAAAGAGAGGCTCCCTGA 1080  
Qy 1082 AAAAGAAAGTTAAATGAACCTTTGAAGTCACACAGGCGCAACTCTTGAAGAAATATATT 1141  
Db 1081 GAAGCGCTTGATGATAGACAGACTGTTGAGAGCCACAGAGTGCATTTGGAATACTCATTT 1140  
Qy 1142 TGCATATTGAAGCAGCAGAGGATTTCTTTAGTGTCAATGCCGATTTTGGCTATAACAGT 1201  
Db 1141 CTGTGCTGCAAGCAATGTTGGGAGCGCAGTGCAGGCTGCTGTAGTGTCTTAGC 1200  
Qy 1202 GTCTTTCTAGCCATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1255  
Db 1201 CAATAATAAACTTTAATGTTGTTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1254

RESULT 14  
AX774279  
LOCUS AX774279 908 bp DNA linear PAT 09-JUL-2003  
DEFINITION Sequence 23 from Patent WO03046176.  
ACCESSION AX774279  
VERSION AX774279.1 GI:32485958  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Legrain, P., Rain, J. C., Benarous, R., Emiliani, S., Berlioz-Torrent, C.  
and Blot, G.  
TITLE Protein-protein interactions in human immunodeficiency virus  
JOURNAL Patent: WO 03046176-A 23 05-JUN-2003;  
Hybrigenics (FR)  
FEATURES  
source  
1. .908  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 70.9%; Score 908; DB 6; Length 908;  
Best Local Similarity 100.0%; Pred. No. 2.9e-184;  
Matches 908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 17 CAAGGAAGGTACCGGCTCTACTGCCACCTCTCCAGCTCCAGCGCGGCGCAGCAGGGA 76

Db 1 CAAGGAAGGTACCGGCTCTACTGCCACCTCTCCAGCTCACCGCGGCGCAGCAGGGA 60  
Qy 77 AGCAAAAGCAAGGCGGCTCGGAGATTCAGCCGTGAAGCAGAGTCAGATAGATGSCCT 136  
Db 61 AGCAAAAGCAAGGCGGCTCGGAGATTCAGCCGTGAAGCAGAGTCAGATAGATGSCCT 120  
Qy 137 TGTGGTATTAAAGATTAATCAACATTTATCAAGAGAGGACCAAGGAATGAAGTTGTTCA 196  
Db 121 TGTGGTATTAAAGATTAATCAACATTTATCAAGAGAGGACCAAGGAATGAAGTTGTTCA 180  
Qy 197 AGGAGTGTCTTTGGGTCTGTTGTAGAAGATCGGCTTTGAAATTTACCAACTGCTTCTCTT 256  
Db 181 AGGAGTGTCTTTGGGTCTGTTGTAGAAGATCGGCTTTGAAATTTACCAACTGCTTCTCTT 240  
Qy 257 CCTCAGCACACAGAGGATGATGCTGAATTTGATGAAGTCCAATATCAGATGGAATGAT 316  
Db 241 CCTCAGCACACAGAGGATGATGCTGATTTGATGAAGTCCAATATCAGATGGAATGAT 300  
Qy 317 GCGGAGCCTTCGCCATGTAAACATTTGATCATCTTCACTGGGCTGGTATCAGTCCACATA 376  
Db 301 GCGGAGCCTTCGCCATGTAAACATTTGATCATCTTCACTGGGCTGGTATCAGTCCACATA 360  
Qy 377 CTATGCTCATTTGCTTTACCCGGGCACCTCTGAGACTCTCAGTTTGTAGTTACAGCATGCCAT 436  
Db 361 CTATGCTCATTTGCTTTACCCGGGCACCTCTGAGACTCTCAGTTTGTAGTTACAGCATGCCAT 420  
Qy 437 TGAAGATCTGTGCTTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCAT 496  
Db 421 TGAAGATCTGTGCTTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCAT 480  
Qy 497 AAAGCATAACAGCTGAGCTCTTAAACTGATGAAGTCTTGTAAAGAAAGGATTTTCCCC 556  
Db 481 AAAGCATAACAGCTGAGCTCTTAAACTGATGAAGTCTTGTAAAGAAAGGATTTTCCCC 540  
Qy 557 TGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAGTCCGATTTGT 616  
Db 541 TGAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAGTCCGATTTGT 600  
Qy 617 AATTAAAAATTCACATCTGATCAATGCTTAATGTTGGGAACTTGAAGAGTACGCTGT 676  
Db 601 AATTAAAAATTCACATCTGATCAATGCTTAATGTTGGGAACTTGAAGAGTACGCTGT 660  
Qy 677 TGCAGATAAATCAATGATTTGCTCAGCTTCCAGCAGCAATCAATTTGGGAAAGTCTTACA 736  
Db 661 TGCAGATAAATCAATGATTTGCTCAGCTTCCAGCAGCAATCAATTTGGGAAAGTCTTACA 720  
Qy 737 GTTGTGATGGACAGAGTGGATGAAATGAGCAAGATATAGTTAAATACACATACAT 796  
Db 721 GTTGTGATGGACAGAGTGGATGAAATGAGCAAGATATAGTTAAATACACATACAT 780  
Qy 797 GAGGAATAGTAAACAAACAGCAGCAGCAATCAGTATCAGCAGCTCGCCAGCAGGA 856  
Db 781 GAGGAATAGTAAACAAACAGCAGCAGCAATCAGTATCAGCAGCTCGCCAGCAGGA 840  
Qy 857 GAATATGAGCGCCAGAGCGAGGAGAACCCCGCTCCCTCAGGAGGACCTGTCCAAACT 916  
Db 841 GAATATGAGCGCCAGAGCGGAGGAGAACCCCGCTCCCTCAGGAGGACCTGTCCAAACT 900  
Qy 917 CTTCAAAAC 924  
Db 901 CTTCAAAAC 908  
RESULT 15  
AC122253/c  
LOCUS AC122253 197161 bp DNA linear ROD 10-JUL-2004  
DEFINITION Mus musculus BAC clone RP23-171K3 from chromosome 17, complete  
sequence.  
ACCESSION AC122253  
VERSION AC122253.5 GI:48675542  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)



```
repeat_region 21396..21601
/rpt_family="MER1_type"
repeat_region 2166R..21768
/rpt_family="B4"
repeat_region 21768..22042
/rpt_family="Alu"
repeat_region 22056..22134
/rpt_family="B4"
repeat_region 22622..22723
/rpt_family="MIR"
repeat_region 23019..23182
/rpt_family="L2"
repeat_region 25183..25287
/rpt_family="B2"
repeat_region 26690..26752
/rpt_family="Alu"
repeat_region 27554..27872
/rpt_family="MaLR"
repeat_region 27924..28312
/rpt_family="L1"
repeat_region 29816..30032
/rpt_family="B2"
repeat_region 30321..30511
/rpt_family="B2"
repeat_region 31586..31666
/rpt_family="Alu"
repeat_region 31779..31821
/rpt_family="Alu"
repeat_region 32756..32851
/rpt_family="AcHobo"
repeat_region 33363..33415
/rpt_family="ERV1"
repeat_region 33891..34204
/rpt_family="B4"
repeat_region 35241..35386
/rpt_family="Alu"
repeat_region 35725..36161
/rpt_family="ERV1"
repeat_region 36841..37155
/rpt_family="MaLR"
repeat_region 37205..37410
/rpt_family="MER1_type"
repeat_region 40718..40915
/rpt_family="B2"
repeat_region 41307..41357
/rpt_family="ERV1"
repeat_region 41426..41477

Query Match 67.8%; Score 868; DB 10; Length 197161;
Best Local Similarity 82.9%; Pred. No. 1.7e-175;
Matches 1065; Conservative 0; Mismatches 205; Indels 14; Gaps 6;

Qy 1 GAAAGATGGCGTCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60
Db 40557 GAAACATGGCGTCCGCAAGGAAGGCACCGCTCCACTGCCACCTCTCCAGTTCTGCTG 40498

Qy 61 CCGCGCAGCAGGGAAGCAAGGCAAGCGGCTCGGAGATTTCAGCGCGTGAAGCAAG 120
Db 40497 GCGGCACGTTGGGGAAGGCAAGGCAAGCGGCTCTGGAGATTCGGTGGTGAAGCAGG 40438

Qy 121 TGCAGATAGATGCGCTGTGGTATTAAAGATATCAACNTATCAAGAAAGAGCAAG 180
Db 40437 TGCAGATCGACGCGCTGTGGTATTAAAGATATCAACNTATCAAGAAAGAGCAAG 40378

Qy 181 GAATCAAGTTGTTCAAGGAGTCTTTTGGGTCGTTGTAGAGATCGGCTTGAAATTA 240
Db 40377 GAATCAAGTTGTTCAAGGAGTCTCTTGGGTCGTTGTAGAGATCGCTAGAGATGA 40318

Qy 241 CCAACTGCTTTCC-TTTCCCTCAGCAACAGAGGATGATGCTTTGATCAAGTCCAA 299
Db 40317 CCAACTGCTTTCCGTTCCCGTCCCGCAGCAGCGGAGGATGATGCTGACTTTTATGAGGTACAG 40258

Qy 300 TATCAGATGGAATATGATGCGGACCTTCGCCATGTAAACATTGATCATCTTCACTGGGC 359
```

```
Db 40257 TATCAGATGGAATATGATGCGCAGCCTTCGCCATGTCAACATTGATCACCTCCACGTGGC 40198
Qy 360 TGGTATCAGTCCACATATACTATGGCTCATTTGTTACCGGGCACTCTCTGGAATCTCAGTTT 419
Db 40197 TGGTATCAGTCCACATATACTATGGCTCTCTGTTACCGGGCGGTTTGGATTTCTCAGTTC 40138
Qy 420 AGTTACAGCATGCCATTTGAAGAAATCTGTGTTCTCATTTATGATCCCATAAAACTGCC 479
Db 40137 AGCTACCAAGCAGCCCATTTGAAGAGTCTGTCTCTCATTTATGATCCCATAAAACTGCC 40078
Qy 480 CAAGGA--TCTCTCTCACTAAAGGCATACA-GACTGACTCTTAAATGATGAGAGTTGT 536
Db 40077 CAAGGATCTCTCTCTCGCTGAAGGCATACATGACTGACTCTTAAACGGATGGAAGTTGT 40018
Qy 537 AAAGAAAAGGATTTTCCCTGAAGCATTTGAAAAAGCAAAATATACACCTTTGAGTACATG 596
Db 40017 AAAGAGAAGGACTTTTCCCTGAAGCATTTGAAAAAGCAAGCATCACCTTTGAGCACATG 39958
Qy 597 TTTGAAGAAAGTCCGATTTGTTAAATTAATAATTCATCTGATCAATGCTCTTAATGTGGAA 656
Db 39957 TTTGAAGAAAGTCCGATTTGTTAAATTAATAATTCATCTGATCAATGCTCTTAATGTGGAG 39898
Qy 657 CTTGAAAAGAAGTCACTGTTGCAGATAAATGAAATGCTCAGCTTCCAGCTTCCAGCAGCAAT 716
Db 39897 CTTGAGAAGAAGTCACTGTTGCGGATTAAGCAGAAATGCTCAGCTTCTAGCAGCAAT 39838
Qy 717 CAT-TTGGGGAAGAATCTACAGTTGCTGATGACAGAGTGGATGAAATGAGCAAGATAT 775
Db 39837 CATCTGGGGAAGAATCTCAGCTGCTGATGACCCGGTGGATGAAATGAGCCAGGACAT 39778
Qy 776 AGTTAAATACACATACATGAGGAATACTAGTAAACAAACAGCAGCAGCAACATCAGTA 835
Db 39777 AATCAAAATACAAACAGCAGCTGCGGAAACACCACTTAACAGCAGCAGCAACATCAGAA 39718
Qy 836 TCAGCAGCGTCCGACAGCAGGAGATATGCAGCGCCAGAGCCAGAGGAGAACCCCTCC 895
Db 39717 TCAGCAGCGCCGCAACAGGAGATATGCAGCGTCGGAGTCGAGCGGAGCCCTCC 39658
Qy 896 TGAGGAGACCTGTCCAACTCTTCAAAACCAACAGCAGCGGCTGCCAGGATGGACTCGT 955
Db 39657 TGAGGAGACCTCTCTAACTCTTCAAGCCCAACAGGCGCTTCCAGGATGGATTCAT 39598
Qy 956 GCTCATTCAGGCCAGATAAACACTTACTGCCAGAACATCAAGGATTCACCTGCCCAAA 1015
Db 39597 GCTCATTCAGGCCAGATGAACACTTCTGCCAGAACATCAAGGATTCACCTGCCCAAA 39538
Qy 1016 CTTAGGCAAGCTCTTTCATGGCCCAAGGCTTTCAAGAAATACAACTAAGAAAAAGGAAGT 1075
Db 39537 CTTAGGCAAACTCTTCATGGCCCAAGGCTTTCAAGAAATACAACTAAGAAAAAGG-AGC 39479
Qy 1076 TCCAGAAAGAGTAAACATGAACCTTTGAAGTCAACACGAGGCAACTCTTGGAGAAA 1135
Db 39478 TTCCTGAGAACGGTTGATGATGACAGACTGTTGAGAGTCCAAGAGTGCACCTTGGAAAA 39419
Qy 1136 TATATTTGCATATTGAAAGCAGCAGAGGATTTCTTTAGTGTCAATTGCCGATTTTGGCTAT 1195
Db 39418 CGTATTTCTGTGCTGCAAGCA-----ATGCTGGGAGCGGCACCTGTGCGTCACTGC 39367
Qy 1196 AACAGTGTCTTTTAGCCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1255
Db 39366 TGTAGTGGTTTTCTAGCCAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39307
Qy 1256 AAAAAAAAAAAAAAAAAAAAAA 1279
Db 39306 AAAAAAAAAAAAAAAAAAAGAGA 39283
```

Search completed: January 5, 2005, 14:32:36  
Job time : 5540 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 12:07:01 ; Search time 4403 Seconds  
(without alignments)  
10593.426 Million cell updates/sec

Title: US-10-017-327-1

Perfect score: 1280

Sequence: 1 gaaagatggctccgcaag.....aaaaaaaaaaaaaaaaaaaa 1280

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1224	95.6	1224	3	CR612538
2	1220	95.3	1222	3	CR611008
3	1181	92.3	1181	3	CR616809
4	1168	91.2	1168	3	CR602203
5	1161	90.7	1165	3	CR590054
6	1093.4	85.4	1102	3	CR597364
7	1028.4	80.3	1158	5	BU902116
8	1006.2	78.6	1122	5	BU507988
9	997	77.9	1076	4	BM551561
c 10	993.8	77.6	1087	1	AL520291
11	984	76.9	1073	4	BM917227
c 12	970.6	75.8	1051	1	AL570102
13	966.6	75.5	1143	1	AL543554
14	966.2	75.5	1014	5	B0228148
15	966	75.5	967	3	CR617128
16	946.4	73.9	999	4	BM905140
17	910.2	71.1	1016	1	AL520292
18	907.2	70.9	985	5	B0675076
19	905.8	70.8	1044	4	BM927525
20	903.4	70.6	930	5	BU500142
21	902.8	70.5	969	5	B0898234
22	900.8	70.4	1028	5	B0062902
23	897.8	70.1	929	5	B0226163
24	897	70.1	1254	3	AK088436

ALIGNMENTS

RESULT 1  
CR612538  
LOCUS CR612538 1224 bp mRNA linear HTC 21-JUL-2004  
DEFINITION full-length cDNA clone CS0DE005YM07 of Placenta of Homo sapiens (human).  
ACCESSION CR612538  
VERSION CR612538.1 GI:50493345  
KEYWORDS HTC; CNSLT CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
2 (bases 1 to 1224)  
REFERENCE  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES  
source  
1..1224  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE005YM07"  
/tissue\_type="Placenta"  
/plasmid="pCMVSPORT\_6"

ORIGIN  
Query Match 95.6%; Score 1224; DB 3; Length 1224;  
Best Local Similarity 100.0%; Pred.No.1:5e-239;  
Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAAAGATGGGTCGCCGAAGGATACCGGCTCTACTGCCACCTCTTCCAGTCCACCG 60  
|||||

Db 1 GAAAGATGGGTCGCCGAAGGATACCGGCTCTACTGCCACCTCTTCCAGTCCACCG 60  
|||||

```
QY 61 CCGGCCGACGAGGGAAGCAAGGCGGCTCGGGAGATTTCAGCCGTTGAAGCAAG 120
Db 61 CCGGCCGACGAGGGAAGCAAGGCGGCTCGGGAGATTTCAGCCGTTGAAGCAAG 120
QY 121 TGCAGATAGATGCGCTTGTGGTATTAAAGATAATCAAACTATTATCAAGAAGAGCAAG 180
Db 121 TGCAGATAGATGCGCTTGTGGTATTAAAGATAATCAAACTATTATCAAGAAGAGCAAG 180
QY 181 GAACTGAAGTTGTTCAAGAGTGCTTTTGGGTCTGGTTGTAGAGATCGGCTTGAATTA 240
Db 181 GAACTGAAGTTGTTCAAGAGTGCTTTTGGGTCTGGTTGTAGAGATCGGCTTGAATTA 240
QY 241 CCAACTGCTTCCCTTCCCTCAGACACAGAGGATGCTGACTTTGATGAAGTCCAAT 300
Db 241 CCAACTGCTTCCCTTCCCTCAGACACAGAGGATGCTGACTTTGATGAAGTCCAAT 300
QY 301 ATCAGATGGAATGATGCGGAGCCTTCCCATGTAACATTTGATCATCTTCACTGGGCT 360
Db 301 ATCAGATGGAATGATGCGGAGCCTTCCCATGTAACATTTGATCATCTTCACTGGGCT 360
QY 361 GGTATCAGTCCACATCTATGCTCATTTGTTACCGGGCACTCTCGACTCTCAGTTTA 420
Db 361 GGTATCAGTCCACATCTATGCTCATTTGTTACCGGGCACTCTCGACTCTCAGTTTA 420
QY 421 GTTACCAGCATGCCATTGAAGAACTCTGTCTCTCATTTATGATCCCATAAACTGCC 480
Db 421 GTTACCAGCATGCCATTGAAGAACTCTGTCTCTCATTTATGATCCCATAAACTGCC 480
QY 481 AAGGATCTCTCTCTAAAGGCATACAGATGACTCTCTAACTGATGGAAGTTTGAAG 540
Db 481 AAGGATCTCTCTCTAAAGGCATACAGATGACTCTCTAACTGATGGAAGTTTGAAG 540
QY 541 AAAAGGATTTTCCCTTGAAGCATTTGAAAAGCAATATACCTTTGATGATCATCTTTG 600
Db 541 AAAAGGATTTTCCCTTGAAGCATTTGAAAAGCAATATACCTTTGATGATCATCTTTG 600
QY 601 AAGAGTGCAGTTGTAATTAATAATTCACATCTGATCAATGCTCTAAATGCGGAATTG 660
Db 601 AAGAGTGCAGTTGTAATTAATAATTCACATCTGATCAATGCTCTAAATGCGGAATTG 660
QY 661 AAAAGAGTGCAGTTGTCAGATAAATGATGATGCTCACTGCTGCGAGCAGCAATCAT 720
Db 661 AAAAGAGTGCAGTTGTCAGATAAATGATGATGCTCACTGCTGCGAGCAGCAATCAT 720
QY 721 TGGGGAAGATCTACAGTTGCTGATGACAGAGTGGATGAAATGAGCAGATATAGTTA 780
Db 721 TGGGGAAGATCTACAGTTGCTGATGACAGAGTGGATGAAATGAGCAGATATAGTTA 780
QY 781 AATACACATACATAGAGGAATCTAGTAAACAAAGCAGCAGAAACATCAGTATCAGC 840
Db 781 AATACACATACATAGAGGAATCTAGTAAACAAAGCAGCAGAAACATCAGTATCAGC 840
QY 841 AGCGTCCGACGAGGAGATATACGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAG 900
Db 841 AGCGTCCGACGAGGAGATATACGCGCCAGAGCCGAGGAGAACCCCGCTCCCTGAG 900
QY 901 AGGACCTGTCCAACTCTTCAACACCACAGCGGCTGCCAGATGAGTCTGCTGCTCA 960
Db 901 AGGACCTGTCCAACTCTTCAACACCACAGCGGCTGCCAGATGAGTCTGCTGCTCA 960
QY 961 TTGAGGCCAGATAAATCACTTACTGCGAGAAATCAAGGAGTTCACTGCCCAAACTTAG 1020
Db 961 TTGAGGCCAGATAAATCACTTACTGCGAGAAATCAAGGAGTTCACTGCCCAAACTTAG 1020
QY 1021 GCAAGCTCTTCATGGCCCGGCTCTTCAAGAAATACAACTAAGAAAGGAGTTTCCA 1080
Db 1021 GCAAGCTCTTCATGGCCCGGCTCTTCAAGAAATACAACTAAGAAAGGAGTTTCCA 1080
QY 1081 GAAAAGAGTTAACTGAAGTCTTTGAAGTCAACAGGCGCACTCTTGTGAAGAAATATAT 1140
Db 1081 GAAAAGAGTTAACTGAAGTCTTTGAAGTCAACAGGCGCACTCTTGTGAAGAAATATAT 1140
```

```
QY 1141 TTGCATATTGAAAAGCACAGAGGATTTCTTAGTGTCAATTGCCGATTTGGCTATACAG 1200
Db 1141 TTGCATATTGAAAAGCACAGAGGATTTCTTAGTGTCAATTGCCGATTTGGCTATACAG 1200
QY 1201 TGTCTTTCTAGCCATAATAATA 1224
Db 1201 TGTCTTTCTAGCCATAATAATA 1224
RESULT 2
CR611008 1222 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0D1081YG20 of Placentia Cot 25-normalized
DEFINITION
ACCESSION CR611008
VERSION 1 GI:50491815
KEYWORDS HTC; CNSLT CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1222)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1222)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
location/Qualifiers
source 1..1222
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081YG20"
/tissue_type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 95.3%; Score 1220; DB 3; Length 1222;
Best Local Similarity 100.0%; Pred. No. 9,7e-239;
Matches 1220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAAAGATGCGCTCCGCAAGGAAGGTACCGCTCTACTGCCACCTCTTCCAGCTCCACG 60
Db 3 GAAAGATGCGCTCCGCAAGGAAGGTACCGCTCTACTGCCACCTCTTCCAGCTCCACG 62
QY 61 CCGCGCAGCAGGGAAGGCAAGGCGGCTCGGAGATTTCAGCCGTTGAAGCAAG 120
Db 63 CCGCGCAGCAGGGAAGGCAAGGCGGCTCGGAGATTTCAGCCGTTGAAGCAAG 122
QY 121 TGCAGATAGATGCGCTTGTGGTATTAAAGATAATCAAACTATTATCAAGAAGAGCAAG 180
Db 123 TGCAGATAGATGCGCTTGTGGTATTAAAGATAATCAAACTATTATCAAGAAGAGCAAG 182
QY 181 GAACTGAAGTTGTTCAAGAGTGCTTTTGGGTCTGGTTGTAGAGATCGGCTTGAATTA 240
Db 183 GAACTGAAGTTGTTCAAGAGTGCTTTTGGGTCTGGTTGTAGAGATCGGCTTGAATTA 242
QY 241 CCAACTGCTTCCCTTCCCTCAGACACAGAGGATGCTGACTTTGATGAAGTCCAAT 300
Db 243 CCAACTGCTTCCCTTCCCTCAGACACAGAGGATGCTGACTTTGATGAAGTCCAAT 302
```

```
Qy 301 ATCAGATCGAAATGATGCGGAGCCTTCGCCATGTAAACATTTGATCATCTTCCACGTGGGCT 360
Db 303 ATCAGATGAAATGATGCGGAGCCTTCGCCATGTAAACATTTGATCATCTTCCACGTGGGCT 362
Qy 361 GGTATCAGTCACATACCTAGCTCATTCGCTTACCGGGCACTCTCGACCTCTCAGTTTA 420
Db 363 GGTATCAGTCACATACCTAGCTCATTCGCTTACCGGGCACTCTCGACCTCTCAGTTTA 422
Qy 421 GTTACCAGATGCCATTAAGAATCTGTCGTTCTCATTTATGATGCCATAAAAACTGCC 480
Db 423 GTTACCAGATGCCATTAAGNATCTGTCCTCTCATTTATGATGCCATAAAAACTGCC 482
Qy 481 AAGGATCTCTCTCACTAAAGCATACAGATGACTCTTAAACTGATGAAAGTTTGTAAAG 540
Db 483 AAGGATCTCTCTCACTAAAGCATACAGATGACTCTTAAACTGATGAAAGTTTGTAAAG 542
Qy 541 AAAAGGATTTTCCCTTGAAGCATTTGAAAAGCAATATCACTTTTTCAGTACATGTTG 600
Db 543 AAAAGGATTTTCCCTTGAAGCATTTGAAAAGCAATATCACTTTTTCAGTACATGTTG 602
Qy 601 AAGAAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGTCTTAATGTGGAACTTG 660
Db 603 AAGAAGTCCGATTTGTAATTAATAATTCACATCTGATCAATGTCTTAATGTGGAACTTG 662
Qy 661 AAAAGAGTCAAGCTTTGTCAGATAAACAATGTAATTTGCTCAGCCTTTCAGCAGCAATCAAT 720
Db 663 AAAAGAGTCAAGCTTTGTCAGATAAACAATGTAATTTGCTCAGCCTTTCAGCAGCAATCAAT 722
Qy 721 TGGGGAAGNATCTACAGTTGCTGATGACAGATGGATGAATGAGCCAAAGATATAGTTA 780
Db 723 TGGGGAAGNATCTACAGTTGCTGATGACAGATGGATGAATGAGCCAAAGATATAGTTA 782
Qy 781 AATCAACACATACATGAGGAATCTAGTAACACAGCAGCAGCAAAACATCAGATATCAGC 840
Db 783 AATCAACACATACATGAGGAATCTAGTAACACAGCAGCAGCAAAACATCAGATATCAGC 842
Qy 841 AGCCTCGCCAGCAGGAGAAATATGACGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 900
Db 843 AGCCTCGCCAGCAGGAGAAATATGACGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 902
Qy 901 AGGACCTGTCAAACCTTTCAAACACACAGCAGCAGCAGCAGATGAGATGCTGTCTCA 960
Db 903 AGGACCTGTCAAACCTTTCAAACACACAGCAGCAGCAGCAGATGAGATGCTGTCTCA 962
Qy 961 TTGCAGGCCAGATAAACAATTTACTTGCAGAGACATCAAGGAGTTCACTGCCCAAACTTAG 1020
Db 963 TTGCAGGCCAGATAAACAATTTACTTGCAGAGACATCAAGGAGTTCACTGCCCAAACTTAG 1022
Qy 1021 GCAAGCTCTTCATGGCCAGGCTCTTCAAGAAATACAACTAAGAAAAGGAAGTTTCCA 1080
Db 1023 GCAAGCTCTTCATGGCCAGGCTCTTCAAGAAATACAACTAAGAAAAGGAAGTTTCCA 1082
Qy 1081 GAAAAGAAAGTTAAATGAACTTTGAAAGTACACAGGGGCAACTTTGGAAGAAATATAT 1140
Db 1083 GAAAAGAAAGTTAAATGAACTTTGAAAGTACACAGGGGCAACTTTGGAAGAAATATAT 1142
Qy 1141 TTGCATATTGAAAAGCACAGAGGATTTCTTTAGTGTCAATTCGCCGATTTTGGCTATAACAG 1200
Db 1143 TTGCATATTGAAAAGCACAGAGGATTTCTTTAGTGTCAATTCGCCGATTTTGGCTATAACAG 1202
Qy 1201 TGTCTTTCTAGCCATAATAA 1220
Db 1203 TGTCTTTCTAGCCATAATAA 1222
```

RESULT 3  
CR616809  
LOCUS CR616809 1181 bp mRNA linear HTC 21-JUL-2004  
DEFINITION full-length cDNA clone CS0D1006YN14 of Placenta Cot 25-normalized  
of Homo sapiens (human).  
ACCESSION CR616809  
VERSION CR616809.1 GI:50497616  
KEYWORDS HTC; CNSLT\_cDNA.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1181)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paraday Avenue  
2 (bases 1 to 1181)  
REFERENCE Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
FEATURES  
source  
1. .1181  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1006YN14"  
/tissue\_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"  
ORIGIN  
Query Match 92.3%; Score 1181; DB 3; Length 1181;  
Best Local Similarity 100.0%; Pred. No. 9e-231;  
Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 20 GGAAGTAGTACCGGCTCTACTGCCACCTCTTCAGAGTCCACCGCCGCGCAGCAGGAAAGG 79  
Db 1 GGAAGTAGTACCGGCTCTACTGCCACCTCTTCAGAGTCCACCGCCGCGCAGCAGGAAAGG 60  
Qy 80 CAAAGGCAAAAGCGGCTCGGAGATTTCAGCGGTGAAGCAAGTGCAGATAGATGCGCTTGT 139  
Db 61 CAAAGGCAAAAGCGGCTCGGAGATTTCAGCGGTGAAGCAAGTGCAGATAGATGCGCTTGT 120  
Qy 140 GGTATTAAAGATATCAAAATATCAAGAAAGGACAAAGAACTGAAGTTGTTCAAG 199  
Db 121 GGTATTAAAGATATCAAAATATCAAGAAAGGACAAAGAACTGAAGTTGTTCAAG 180  
Qy 200 AGTGTCTTTGGGTCTGGTCTGTAGAAAGATCGGCTTTGAAATACCAACTGCTTTCTCTTCCC 259  
Db 181 AGTGTCTTTGGGTCTGGTCTGTAGAAAGATCGGCTTTGAAATACCAACTGCTTTCTCTTCCC 240  
Qy 260 TCAGCACACAGAGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATATGATGCG 319  
Db 241 TCAGCACACAGAGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATATGATGCG 300  
Qy 320 GAGCCTTCGCGCATGTAACATTTGATCATCTTCAGTGGGCTGGTATCAGTCCACATCTA 379  
Db 301 GAGCCTTCGCGCATGTAACATTTGATCATCTTCAGTGGGCTGGTATCAGTCCACATCTA 360  
Qy 380 TGGCTCATCTGTTACCGGGCACTCTCGGACTCTCAGTCTCAGTTAGTTACCAGATGCCATTGA 439  
Db 361 TGGCTCATCTGTTACCGGGCACTCTCGGACTCTCAGTCTCAGTTAGTTACCAGATGCCATTGA 420  
Qy 440 AGAATCTGTCTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCAGCTAAA 499  
Db 421 AGAATCTGTCTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCAGCTAAA 480  
Qy 500 GGATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTCTGA 559  
Db 481 GGATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTCTGA 540  
Qy 560 AGCATTGAAAAAGCAAAATATACCTTTTGAGTACATGTTTGAAGAGAGTGGCGGATTGTAAT 619



```
Db 541 AGCAATGAAAAAGCAATATACCTTGGAGTCAATGTTGAAGAGTGGCGATTGTAAT 600
Qy 620 TAAATAATTCACATCTGATCAATGTCCTTAATGFGGAACTTGAAAAGAAAGTCAGCTGTGC 679
Db 601 TAAATAATTCACATCTGATCAATGTCCTTAATGFGGAACTTGAAAAGAAAGTCAGCTGTGC 660
Qy 680 AGATAAACATGAATTTGCTCAGCCTTGGCAGCAGCAATCATTTGGGGAAGAAATCTACAGTT 739
Db 661 AGATAAACATGAATTTGCTCAGCCTTGGCAGCAGCAATCATTTGGGGAAGAAATCTACAGTT 720
Qy 740 GCTGATGACAGAGTGGATGAATAGAGCAAGATATAGTTAAATACACATACATGAG 799
Db 721 GCTGATGACAGAGTGGATGAATAGAGCAAGATATAGTTAAATACACATACATGAG 780
Qy 800 GAATACTAGTAAACAAACAGCAGCAAGAAATCATGATATCAGAGCGTGCAGCAGGAGAA 859
Db 781 GAATACTAGTAAACAAACAGCAGCAAGAAATCATGATATCAGAGCGTGCAGCAGGAGAA 840
Qy 860 TATGAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTCCAACTCTT 919
Db 841 TATGAGCGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTCCAACTCTT 900
Qy 920 CAAACCAACACAGCGCCTGCGCAGGATGGAATCGCTCATTTGAGGAGGACCTGTCCAACTCTT 979
Db 901 CAAACCAACACAGCGCCTGCGCAGGATGGAATCGCTCATTTGAGGAGGACCTGTCCAACTCTT 960
Qy 980 TTACTGCCAGAACATCAAGGAGTTCACCTGCCCAAACTTAGGCAAGCTCTTCAATGAGG 1039
Db 961 TTACTGCCAGAACATCAAGGAGTTCACCTGCCCAAACTTAGGCAAGCTCTTCAATGAGG 1020
Qy 1040 GGCTCTTCAAGAAATACAACTAAGAAAGAAAGTTCAGAAAGAAAGTTCAGAAAGTTCAGAA 1099
Db 1021 GGCTCTTCAAGAAATACAACTAAGAAAGAAAGTTCAGAAAGAAAGTTCAGAAAGTTCAGAA 1080
Qy 1100 CTCTTGAAGTACACAGGAGCACTCTTGAAGAAATATATTTGCAATATGTAAGAAAGCA 1159
Db 1081 CTCTTGAAGTACACAGGAGCACTCTTGAAGAAATATATTTGCAATATGTAAGAAAGCA 1140
Qy 1160 GAGGATTTCTTTAGTGTTCATTGCGGATTTGGCTATTAACAG 1200
Db 1141 GAGGATTTCTTTAGTGTTCATTGCGGATTTGGCTATTAACAG 1181

RESULT 4
CR602203
LOCUS full-length cDNA clone CS0DA003YL13 of Neuroblastoma of Homo sapiens (human).
DEFINITION CR602203 1168 bp mRNA linear HTC 21-JUL-2004
ACCESSION CR602203
VERSION CR602203.1 GI:50483010
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1168)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1168)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]
COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime ends enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a
```

```
division of Invitrogen.
FEATURES
source 1. .1168
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA003YL13"
/tissue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 91.2%; Score 1168; DB 3; Length 1168;
Best Local Similarity 100.0%; Pred. No. 4.1e-228;
Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 15 CGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGCGCCGACGAGG 74
Db 1 CGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGCGCGACGAGG 60
Qy 75 AAAGGCAAAAGGCAAAAGGCGGCTCGGAGATTGAGCCGTGAAGCAAGTCAGATAGTGC 134
Db 61 AAAGGCAAAAGGCAAAAGGCGGCTCGGAGATTGAGCCGTGAAGCAAGTCAGATAGTGC 120
Qy 135 CTGTGTGATTAAAGATATCAACATTTATCAAGAGAGGCAAGGAACTGAAGTTGTT 194
Db 121 CTGTGTGATTAAAGATATCAACATTTATCAAGAGAGGCAAGGAACTGAAGTTGTT 180
Qy 195 CAAGGAGTGCTTTTGGGCTGTGTTGTAAGATCGGCTTGAAATTAACCACTGCTTCT 254
Db 181 CAAGGAGTGCTTTTGGGCTGTGTTGTAAGATCGGCTTGAAATTAACCACTGCTTCT 240
Qy 255 TTCCCTCAGCACACAGAGGATGCTGACTTTGATGAAGTCCAAATATCAGATGAAATG 314
Db 241 TTCCCTCAGCACACAGAGGATGCTGACTTTGATGAAGTCCAAATATCAGATGAAATG 300
Qy 315 ATGCGGAGCCTTCGGCATGTAAACATTCATCATCTTACGTGGGTGCTATCAGTCCACA 374
Db 301 ATGCGGAGCCTTCGGCATGTAAACATTCATCATCTTACGTGGGTGCTATCAGTCCACA 360
Qy 375 TACTATGGCTCATTCGTTTACCGGGCACTCTCGGACTCTCAGTTTAGTTACGAGCATGCC 434
Db 361 TACTATGGCTCATTCGTTTACCGGGCACTCTCGGACTCTCAGTTTAGTTACGAGCATGCC 420
Qy 435 ATTGAAGAATCTGCTGTTCTCATTTATGATCCCAATATAAACTGCCAAGGATCTCTCA 494
Db 421 ATTGAAGAATCTGCTGTTCTCATTTATGATCCCAATATAAACTGCCAAGGATCTCTCA 480
Qy 495 CTAAGGCAATACAGACTGACTCTTAACTGATGGAAGTTTGTAAAGAAAAGGATTTTCC 554
Db 481 CTAAGGCAATACAGACTGACTCTTAACTGATGGAAGTTTGTAAAGAAAAGGATTTTCC 540
Qy 555 CCTGAGCATTTGAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAAAGTCCGANT 614
Db 541 CCTGAGCATTTGAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAAAGTCCGANT 600
Qy 615 GTAAATAAAAATTCACATCTGATCAATGTCCTAATGTGGGAACCTTGAAAAGAAAGTCAGT 674
Db 601 GTAAATAAAAATTCACATCTGATCAATGTCCTAATGTGGGAACCTTGAAAAGAAAGTCAGT 660
Qy 675 GTTCAGATAAACATGAATTTGCTCAGCCTTGCAGCAGCAATCATTTGGGGAAGAAATCTA 734
Db 661 GTTCAGATAAACATGAATTTGCTCAGCCTTGCAGCAGCAATCATTTGGGGAAGAAATCTA 720
Qy 735 CAGTTGCTGATGGACAGAGTGAATGAAATGAGCCAGATATAGTTAAATACACATATAC 794
Db 721 CAGTTGCTGATGGACAGAGTGAATGAAATGAGCCAGATATAGTTAAATACACATATAC 780
Qy 795 ATGAGGAATACTAGTAAACAAACAGCAGCAAGAAATCATCAGTATCAGCAGCGTCGCGACAG 854
Db 781 ATGAGGAATACTAGTAAACAAACAGCAGCAAGAAATCATCAGTATCAGCAGCGTCGCGACAG 840
Qy 855 GAGAAATGAGCGCCAGCAGGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTGTCAAA 914
Db 841 GAGAAATGAGCGCCAGCAGGAGGAGGAGAAACCCCGCTCCCTGAGGAGGACCTGTGTCAAA 900
```



Db 841 GAGAAATATGACGCCAGAGCCGAGGAGAAACCCCGCTCCCTGAGGAGGACCTCTCCAAA 900  
Qy 915 CTCTTTAAACACCAACAGCCGCTGCCAGGATGACCTGCTGCTCATTTGACAGGCGCAGATA 974  
Db 901 CTCTTTAAACACCAACAGCCGCTGCCAGGATGACCTGCTGCTCATTTGACAGGCGCAGATA 960  
Qy 975 AACACTTACTGCCAGAACATCAAGAGTTCTACCTGCCAAACCTTAGGCAAGCTCTTCATG 1034  
Db 961 AACACTTACTGCCAGAACATCAAGAGTTCTACCTGCCAAACCTTAGGCAAGCTCTTCATG 1020  
Qy 1035 GCCCAGGCTCTTCAAGATACAACTAGAAAGGAGTTTCCAGAAAGAGTTAAC 1094  
Db 1021 GCCCAGGCTCTTCAAGATACAACTAGAAAGGAGTTTCCAGAAAGAGTTAAC 1080  
Qy 1095 ATGAACCTTTGAACTCACACAGGCGCAACTCTTGGAGAAATATATTTGTCATATTGAAA 1154  
Db 1081 ATGAACCTTTGAACTCACACAGGCGCAACTCTTGGAGAAATATATTTGTCATATTGAAA 1140  
Qy 1155 GCACAGAGGATTTCTTAGTGTCATTGC 1182  
Db 1141 GCACAGAGGATTTCTTAGTGTCATTGC 1168

RESULT 5  
CR590054  
LOCUS full-length cDNA clone CS0DB09YJ15 of Neuroblastoma Cot 1165 bp mRNA linear HTC 21-JUL-2004  
DEFINITION 10-normalized of Homo sapiens (human).  
ACCESSION CR590054  
VERSION CR590054.1 GI:50470709  
KEYWORDS HTC; CNSLT\_cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1165)  
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Feng Liang Email: fliang@lifetech.com URL:  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
Genoscope.  
2 (bases 1 to 1165)  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
Location/Qualifiers  
1. .1165  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DB09YJ15"  
/tissue\_types="Neuroblastoma Cot 10-normalized"  
/plasmid="pCMVSPORT\_6"

ORIGIN  
Query Match 90.7%; Score 1161; DB 3; Length 1165;  
Best Local Similarity 100.0%; Pred. No. 1.le-226;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GAAAGATGGGCTCCCGCAAGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Db 5 GAAAGATGGGCTCCCGCAAGAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCG 64  
Qy 61 CCGCGCAGCAGGGAAGGCAAGGCAAGCGGCTCGGAGATTTCAGCGTGAAGCAAG 120  
|||||

Db 65 CCGCGCAGCAGGGAAGGCAAGGCAAGCGGCTCGGAGATTTCAGCCGCTGAAGCAAG 124  
Qy 121 TGCAGATAGATGGCTTTGTGTTATTAAGATAATCAAAACATTAACAAGAAAGCAAG 180  
Db 125 TGCAGATAGATGGCTTTGTGTTATTAAGATAATCAAAACATTAACAAGAAAGCAAG 184  
Qy 181 GAACCTGAAGTTGTTCAAGGAGTGTCTTTTGGTCTGGTTGTAGAAAGATCGGCTTGAAATTA 240  
Db 185 GAACCTGAAGTTGTTCAAGGAGTGTCTTTTGGTCTGGTTGTAGAAAGATCGGCTTGAAATTA 244  
Qy 241 CCAACTGCTTTTCCCTTCCCTCAGCAGCAGAGGATGATGCTGCTTTTGTATGAAGTCCAAT 300  
Db 245 CCAACTGCTTTTCCCTTCCCTCAGCAGCAGAGGATGATGCTGCTTTTGTATGAAGTCCAAT 304  
Qy 301 ATCAGATGGAATATGATCGGAGCCTTCGCCATGTAAACATTTGATCATCTTTCACCTGGGCT 360  
Db 305 ATCAGATGGAATATGATCGGAGCCTTCGCCATGTAAACATTTGATCATCTTTCACCTGGGCT 364  
Qy 361 GGTATCAGTCCACATATCTATGGCTCATTCGTTACCCGGGCACTCTCTGGACTCTCAGTTTA 420  
Db 365 GGTATCAGTCCACATATCTATGGCTCATTCGTTACCCGGGCACTCTCTGGACTCTCAGTTTA 424  
Qy 421 GTTACCAGCATGCCATTTGAAGAATCTGCTGTTCTCATTTATGATCCCATAAAACTGCC 480  
Db 425 GTTACCAGCATGCCATTTGAAGAATCTGCTGTTCTCATTTATGATCCCATAAAACTGCC 484  
Qy 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACATGATGAAGTTTGTAAAG 540  
Db 485 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAAACATGATGAAGTTTGTAAAG 544  
Qy 541 AAAAGGATTTTCCCTTGAAGCATTTGAAAAAGCAAAATATCACCTTTTGAGTACATGTTT 600  
Db 545 AAAAGGATTTTCCCTTGAAGCATTTGAAAAAGCAAAATATCACCTTTTGAGTACATGTTT 604  
Qy 601 AAGAGTGGCGATGTAAATTTAAAAATTTCACTCTGATCAATGCTCTTAATGTGGAACTTG 660  
Db 605 AAGAGTGGCGATGTAAATTTAAAAATTTCACTCTGATCAATGCTCTTAATGTGGAACTTG 664  
Qy 661 AAAAGAGTCAAGCTGTTGCGAGATAAACATGATGCTCAGCTTGCAGCAGCAGCAATCAT 720  
Db 665 AAAAGAGTCAAGCTGTTGCGAGATAAACATGATGCTCAGCTTGCAGCAGCAGCAATCAT 724  
Qy 721 TGGGGAAGAAATCTACAGTTCTCTGATGGAACAGAGTGGATGAAATGAGCCAAAGATATGTTA 780  
Db 725 TGGGGAAGAAATCTACAGTTCTCTGATGGAACAGAGTGGATGAAATGAGCCAAAGATATGTTA 784  
Qy 781 AATACACATACATACATGAGGAATCTAGTAAACACAGCAGCAGCAATCAGTATCAGC 840  
Db 785 AATACACATACATGAGGAATCTAGTAAACACACAGCAGCAGCAATCAGTATCAGC 844  
Qy 841 AGCTGCCAGCAGGAGGAATATGAGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 900  
Db 845 AGCTGCCAGCAGGAGGAATATGAGCGCCAGAGCCGAGAGAACCCCGCTCCCTGAGG 904  
Qy 901 AGGACCTGTCCAAACTCTTCAAAACCCACACAGCGCGCTGCCAGGATGGACTCGCTGTCTCA 960  
Db 905 AGGACCTGTCCAAACTCTTCAAAACCCACACAGCGCGCTGCCAGGATGGACTCGCTGTCTCA 964  
Qy 961 TTGCAGGCCAGATAAACAATTACTTCCAGAACTCAAGGAGTTTCACTGCCCAAACTTAG 1020  
Db 965 TTGCAGGCCAGATAAACAATTACTTCCAGAACTCAAGGAGTTTCACTGCCCAAACTTAG 1024  
Qy 1021 GCAAGCTCTTCATGGCCCGGCTCTTCAAGAATAACAACATTAAGAAAGGAGTTTCCA 1080  
Db 1025 GCAAGCTCTTCATGGCCCGGCTCTTCAAGAATAACAACATTAAGAAAGGAGTTTCCA 1084  
Qy 1081 GAAAAGAGTTTAAACATGAACCTTTTGAAGTCAACAGGCGCAACTCTTTGGAAGAAATATAT 1140  
Db 1085 GAAAAGAGTTTAAACATGAACCTTTTGAAGTCAACAGGCGCAACTCTTTGGAAGAAATATAT 1144  
Qy 1141 TTGCATATTGAAAAGCAGACA 1161  
Db 1145 TTGCATATTGAAAAGCAGACA 1165

```
RESULT 6
CRS97364
LOCUS
DEFINITION
full-length cDNA clone CS0DB006YK07 of Neuroblastoma Cot
10-normalized of Homo sapiens (human).
ACCESSION
CRS97364
VERSION
1
KEYWORDS
HTC; CNSLT_cDNA..
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1102)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1102)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1102
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB006YK07"
/tissue_type="Neuroblastoma Cot 10-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 85.4%; Score 1093.4; DB 3; Length 1102;
Best Local Similarity 99.9%; Pred. No. 7e-213;
Matches 1094; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 138 GTGGTATTAAAGTATCAACATTATCAAGAGAGGACAGGAACTGAAGTTGTTCAA 197
DB 1 GTGGTATTAAAGTATCAACATTATCAAGAGAGGACAGGAACTGAAGTTGTTCAA 60
QY 198 GGAGTGTCTTTGGGTCTGGTGTAGAGAGATCGGCTTGAATTAACCACTGCTTCTCTTC 257
DB 61 GGAGTGTCTTTGGGTCTGGTGTAGAGAGATCGGCTTGAATTAACCACTGCTTCTCTTC 120
QY 258 CTTACGACACACAGAGATGATGCTGATTTGATGAAGTCCAAATATCAGATGGAATGATG 317
DB 121 CTTACGACACACAGAGATGATGCTGATTTGATGAAGTCCAAATATCAGATGGAATGATG 180
QY 318 CGGAGCTTCGCCATGTAACATTGATCATCTTCAGTGGGTGGTATCAGTCCACATAC 377
DB 181 CGGAGCTTCGCCATGTAACATTGATCATCTTCAGTGGGTGGTATCAGTCCACATAC 240
QY 378 TATGGTCTATTGCTTACCCGGGCACCTCTCGGACTCTCAGTTAGTTTACCAGATGCCATT 437
DB 241 TATGGTCTATTGTTACCCGGGCACCTCTCGGACTCTCAGTTAGTTTACCAGATGCCATT 300
QY 438 GAAGATCTGTCGTTCTCTATTATGATCCATAAAACCTGCCCAAGGATCTCTCTCACTA 497
DB 301 GAAGATCTGTCGTTCTCTATTATGATCCATAAAACCTGCCCAAGGATCTCTCTCACTA 360
QY 498 AAGGCATACAGACTGACTCTCTAACTGATGGAGTTTGTAAAGAAAGATTTTCCCT 557
DB 361 AAGGCATACAGACTGACTCTCTAACTGATGGAGTTTGTAAAGAAAGATTTTCCCT 420
```

```
QY 558 GAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAAAGTGGGATTGTA 617
DB 421 GAAGCATTTGAAAAAGCAAAATATCACCTTTGAGTACATGTTTGAAGAAAGTGGGATTGTA 480
QY 618 ATTAAAAATTACATCTGATCAATGTCCTAAATGTGGAACTTTGAAAGAAAGTCAAGCTGTT 677
DB 481 ATTAAAAATTACATCTGATCAATGTCCTAAATGTGGAACTTTGAAAGAAAGTCAAGCTGTT 540
QY 678 GCAGATAAACATGAATTTGCTCAGCCTTGCACGACAGCAATCATTTGGGGAAAGATCTACAG 737
DB 541 GCAGATAAACATGAATTTGCTCAGCCTTGCACGACAGCAATCATTTGGGGAAAGATCTACAG 600
QY 738 TTGCTGATGGACAGAGTGGATGAATGAGCCCAAGATAGTTAAATACACACATACATG 797
DB 601 TTGCTGATGGACAGAGTGGATGAATGAGCCCAAGATAGTTAAATACACACATACATG 660
QY 798 AGGAATACCTAGTAAAAACAACAGCAGCAGAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAG 857
DB 661 AGGAATACCTAGTAAAAACAACAGCAGCAGAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAG 720
QY 858 AATATGCGCGCCAGCGAGGAGAGAACCCCGCTCCCTGAGGAGGACCTGTCCAAACTC 917
DB 721 AATATGCGCGCCAGCGAGGAGAGAACCCCGCTCCCTGAGGAGGACCTGTCCAAACTC 780
QY 918 TTCAACACACACACAGCGCTGCCAGGATGGAATCGCTGCTCATTTGCGAGGCGCAGATTAAC 977
DB 781 TTCAACACACACACAGCGCTGCCAGGATGGAATCGCTGCTCATTTGCGAGGCGCAGATTAAC 840
QY 978 ACTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAAACTTTAGGCAAGCTCTTCATGGCC 1037
DB 841 ACTTACTGCCAGAAACATCAAGGAGTTCACTGCCCAAAACTTTAGGCAAGCTCTTCATGGCC 900
QY 1038 CAGGCTCTTCAAGATAACAACAATAAGAAAGGAGTTTCCAGAAAGAAAGTAAACATG 1097
DB 901 CAGGCTCTTCAAGATAACAACAATAAGAAAGGAGTTTCCAGAAAGAAAGTAAACATG 960
QY 1098 AACTCTTGAAGTCACACAGGCGCACTCTCGAGAGAAATATATTGCAATATTGAAAGCA 1157
DB 961 AACTCTTGAAGTCACACAGGCGCACTCTCGAGAGAAATATATTGCAATATTGAAAGCA 1020
QY 1158 CAGAGGATTTCTTTAGTGTCTATTGCGGATTTGGCTATAACAGTGTCTTTCTAGCCATAA 1217
DB 1021 CAGAGGATTTCTTTAGTGTCTATTGCGGATTTGGCTATAACAGTGTCTTTCTAGCCATAA 1080
QY 1218 TAAAAATAAAAAA 1232
DB 1081 TAAAAATAAAAAA 1095
RESULT 7
BU902116
LOCUS
DEFINITION
AGENCOURT_10127389 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502779
5', mRNA sequence.
ACCESSION
BU902116
VERSION
BU902116.1 GI:24084029
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1158)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM14059 row: g column: 04

High quality sequence stop: 818.

#### FEATURES

source  
1. .1158  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:6502779"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: NotI;  
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb. "

#### ORIGIN

Query Match 80.3%; Score 1028.4; DB 5; Length 1158;  
Best Local Similarity 96.0%; Pred. No. 1.3e-199;  
Matches 1097; Conservative 0; Mismatches 42; Indels 4; Gaps 4;  
  
Qy 1 GAAAGATGGCTCCGCAAGGAAGGTACCGCTCTACTGCCACCTCTTCCAGCTCCACCG 60  
Db 14 GAAAGATGGCTCCGCAAGGAAGGTACCGCTCTACTGCCACCTCTTCCAGCTCCACCG 73  
  
Qy 61 CCGCGCAGCAGGGAAGGCAAGGCAAGGCGCTCGGAGATTTCAGCCGTGAAGCAAG 120  
Db 74 CCGCGCAGCAGGGAAGGCAAGGCGCTCGGAGATTTCAGCCGTGAAGCAAG 133  
  
Qy 121 TGCAGATAGATGGCTTGTGTATTAAGATAATCAAAACATTATCAAGAAAGGCAAG 180  
Db 134 TGCAGATAGATGGCTTGTGTATTAAGATAATCAAAACATTATCAAGAAAGGCAAG 193  
  
Qy 181 GAATGAAGTTGTCAGGAGTGTCTTTGGTCTGGTTGTAGAAGATCGGCTTCAATTA 240  
Db 194 GAATGAAGTTGTCAGGAGTGTCTTTGGTCTGGTTGTAGAAGATCGGCTTCAATTA 253  
  
Qy 241 CCAACTGCTTTCCCTCCCTCAGCACACAGAGGATGCTGACTTTTGATGAAGTCCAAT 300  
Db 254 CCAACTGCTTTCCCTCCCTCAGCACACAGAGGATGCTGACTTTTGATGAAGTCCAAT 313  
  
Qy 301 ATCAGATGGAATGATCGGAGCTTCGCCATGTAACATTGATCATCTTCACTGGGCT 360  
Db 314 ATCAGATGGAATGATCGGAGCTTCGCCATGTAACATTGATCATCTTCACTGGGCT 373  
  
Qy 361 GGTATCAGTCCACATATAGGCTCATTCGTACCCGGGCACTCTCGACTCTCAGTTTA 420  
Db 374 GGTATCAGTCCACATATAGGCTCATTCGTACCCGGGCACTCTCGACTCTCAGTTTA 433  
  
Qy 421 GTTACACAGATGCCATTGAAGAACTCTGCTCTCTATTATGATCCCAATAAAAGTCCC 480  
Db 434 GTTACACAGATGCCATTGAAGAACTCTGCTCTCTATTATGATCCCAATAAAAGTCCC 493  
  
Qy 481 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGAAG 540  
Db 494 AAGGATCTCTCTCACTAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGAAG 553  
  
Qy 541 AAAAGGATTTTCCCTCGAAGCATTTGAAAAAGCAATATCATCCTTTGAGTACATGTTG 600  
Db 554 AAAAGGATTTTCCCTCGAAGCATTTGAAAAAGCAATATCATCCTTTGAGTACATGTTG 613  
  
Qy 601 AAGAGTGGCGATTGTAATTAATAATTCACATCTGATCAATGCTTAATGTGGGAAGTTG 660  
Db 614 AAGAGTGGCGATTGTAATTAATAATTCACATCTGATCAATGCTTAATGTGGGAAGTTG 673  
  
Qy 661 AAAAGAGTGCAGTGTGTCAGATAAATCAATGAATGCTCAGCTTGCAGCAGCAATCAT 720  
Db 674 AAAAGAGTGCAGTGTGTCAGATAAATCAATGAATGCTCAGCTTGCAGCAGCAATCAT 733  
  
Qy 721 TGGGGAAGATCTACAGTTGCTGATGGACAGATGGATGAAATGAGGCAAGATATAGTTA 780  
Db 734 TGGGGAAGATCTACAGTTGCTGATGGACAGATGGATGAAATGAGGCAAGATATAGTTA 793

Qy 781 AATACACACATACATCAGGAATAGTAGTAAACACAGCAGCAGAAACATCAGTATCAGC 840  
Db 794 AATACACACATACATCAGGAATAGTAGTAAACACAGCAGCAGAAACATCAGTATCAGC 853  
  
Qy 841 AGCGTCGCCAGCAGGAGAAATATGCGCGCCAGAGCCGAGGAGAACCCCGCTCCTCAGG 900  
Db 854 AGCGTCGCCAGCAGGAGAAATATGCGCGCCAGAGCCGAGGAGAACCCCGCTCCTCAGG 913  
  
Qy 901 AGACCTGTCTCAAACTCTTCAAAACACACAGCGCGCTGCCAGGATGGACTCGCTGTCTCA 960  
Db 914 AGACCTGTCTCAAACTCTTCAAA-CACACAGCGCGCTGCCAGGATGGACTCGCTGTCTCA 972  
  
Qy 961 TTGCAGGCCAGATAAACAATCTTACTGCCAGAACATCAAGGAGTTTCACTGCCCAAACTTAG 1020  
Db 973 TGCAGGCCAGATAAACAATCTTACTGCC-GAACATCAAGGAGTTTCCCTGGCCCAAACTAG 1031  
  
Qy 1021 GCAAGCTCTTCATGGCCCGAGGCTCTTCAAGAAATACAACTAAGAAAGGAGTTTCCA 1080  
Db 1032 GCAAGCTCTTCCTGGCCCGAGGCTTCTTCAGAAAT-CCACACCTTAGAAAAAGGAAATTTCCG 1090  
  
Qy 1081 GAAAAGAAAGTTAAACATGAAC-TCTTGAAGTCACACAGGCGCAACTCTTGGAGAAATATA 1139  
Db 1091 AAAAAAGAAATTTACTTGAAGCTTCTTGAAGCCCCCGGCGCAATTTTGGAAAAAATATT 1150  
  
Qy 1140 TTT 1142  
Db 1151 TTT 1153  
  
RESULT 8  
BU507988  
LOCUS BU507988 1122 bp mRNA linear EST 12-SEP-2002  
DEFINITION AGENCOURT 10127980 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6502012  
5', mRNA sequence.  
ACCESSION BU507988  
VERSION BU507988.1 GI:22814221  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1122)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM14057 row: g column: 05  
High quality sequence stop: 779.  
FEATURES  
source  
1. .1122  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:6502012"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: NotI;  
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb. "

Query Match 78.6%; Score 1006.2; DB 5; Length 1122;  
Best Local Similarity 95.1%; Pred. No. 4.6e-195;



```
Db 241 TCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAATGA 300
Qy TCGGAGGCTTCGCCATGTAACATTTGATCATCTTTCAGTGGCTGGTATCAGTCCACAT 375
Db 301 TCGGAGGCTTCGCCATGTAACATTTGATCATCTTTCAGTGGCTGGTATCAGTCCACAT 360
Qy 376 ACTATGGCTCAATTCGTTACCCGGGCACCTCTGAGCTCTCAGTTTATGTTACCAAGCATGCCA 435
Db 361 ACTATGGCTCAATTCGTTACCCGGGCACCTCTGAGCTCTCAGTTTATGTTACCAAGCATGCCA 420
Qy 436 TTGAAGAATCTGCTGTTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCAC 495
Db 421 TTGAAGAATCTGCTGTTCTCATTTATGATCCCATAAAACTGCCCAAGGATCTCTCTCAC 480
Qy 496 TAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCC 555
Db 481 TAAAGGCATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCC 540
Qy 556 CTGAAGCATTTGAAAAAGCAATATCACCTTTGAGTACATGTTTGAAGAAGTGCAGATTG 615
Db 541 CTGAAGCATTTGAAAAAGCAATATCACCTTTGAGTACATGTTTGAAGAAGTGCAGATTG 600
Qy 616 TAAATAAAAATTCACATCTGATCAATGTCCTTAATGTGGGAATCTGAAAAAGATCAGCTG 675
Db 601 TAAATAAAAATTCACATCTGATCAATGTCCTTAATGTGGGAATCTGAAAAAGATCAGCTG 660
Qy 676 TTGCAGATAAACATGAATTTGCTCAGCTTGCAGCAGCAATCATTTGGGGAAGATCTTAC 735
Db 661 TTGCAGATAAACATGAATTTGCTCAGCTTGCAGCAGCAATCATTTGGGGAAGATCTTAC 720
Qy 736 AGTTGCTGATGGACAGAGTGGATGAATGAGCAGAGATATAGTTAAATACACACATACA 795
Db 721 AGTTGCTGATGGACAGAGTGGATGAATGAGCAGAGATATAGTTAAATACACACATACA 780
Qy 796 TGAGGAATCTAGTAAACAAACAGCAGAGAGAAACATCAGTATCAGCAGCGTGCAGCAGG 855
Db 781 TGAGGGAATCTAGTAAACAAACAGCAGAGAGAAACATCAGTATCAGCAGCGTGCAGCAGG 840
Qy 856 AGAATATCAGCGCCAGAGCGGAGAGAACCCCGCT-CCCTGAGGAGGACTGTCCAAA 914
Db 841 AGAATATCAGCGCCAGAGCGGAGAGAACCCCGCTCCCTCGAGGAGGACTGTCCAAA 900
Qy 915 CTCTTCAACACACACAGCGCTGCCAGGATGAGTCTGCTCATTTGAGGCGCAGATA 974
Db 901 CTCCTTCAACACACAGCGCTGCCAGGATGAGTCTGCTCATTTGAGGCGCAGATA 960
Qy 975 AACACTTACTGCGAGAAACATCAAGAGATTCTAGTCCCAAACTTAGGCAAGCTCTTCATG 1034
Db 961 AAACCTTACTGCGAAAACTTCAAGAAATTCTGCGCCCAAACTTAGGCAAGCTTTTTCATG 1020
Qy 1035 GCCAGGCTCTTCAAGATACACAACTAAGAAAGGAGTTTCCAGAAAGAGTT 1091
Db 1021 GCCAGGCTCTTCAAGATAC-ACCCCTAAGAAAGGAGTTTCCAAAGAGTT 1076
```

```
RESULT 10
AL520291/c
LOCUS AL520291 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DB006YK07 3-PRIME, mRNA sequence.
ACCESSION AL520291
VERSION AL520291.3 GI:45695831
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1087)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31038632.
```

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4657.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0DB006AF04NP1&c=4657.r.

FEATURES

source

Location/Qualifiers  
1..1087  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DB006YK07"  
/tissue\_type="NEUROBLASTOMA COT 10-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 77.6%; Score 993.8; DB 1; Length 1087;  
Best Local Similarity 95.1%; Pred. No. 1.6e-192;  
Matches 1024; Conservative 10; Mismatches 42; Indels 1; Gaps 1;  
Qy 156 AACATTATCAAGAAGAGGACAAAGTCTGAAGTTGTTCAAGGAGTGCTTTTGGGTCTG 215  
Db 1083 AARWTATCAACATWTCCAGGAGAGACARGACTGNAKTGTTCAAGAGTGCTTTGGTCT 1024  
Qy 216 GTTGTAGAGATCGGCTTGAATTAACCACTGCTTCTCTTCCCTCAGCAGCAGAGAT 275  
Db 1023 GTTGTAGAAGA-CGCGTGAATTAACCACTGCTTCTCTCTCCCTCAGCAGCAGAGAT 965  
Qy 276 GATGCTGACTTTGATGAAGTCCAATATCAGATGAAATGATGCGGAGCTTCGCCATGTA 335  
Db 964 GATGCTGACTTTGATGAAGTCCAATATCAGATGAAATGATGCGGAGCTTCGCCATGTA 905  
Qy 336 AACATTGATCATCTTACGCTGGCTGGTATCAGTCCACATATATGGCTCATTCGTTACC 395  
Db 904 AACATTGATCATCTTACGCTGGCTGGTATCAGTCCACATATATGGCTCATTCGTTACC 845  
Qy 396 CGGCACTCTCGACTCTCAGTTTATGTTACAGATGCGCATGAAGATCTGTCTCTC 455  
Db 844 CGGCACTCTCTGAGCTCTCAGTTTATGTTACAGATGCGCATGAAGATCTGTCTCTC 785  
Qy 456 ATTTATGATCCATAAAAACTGCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACT 515  
Db 784 ATTTATGATCCATAAAAACTGCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACT 725  
Qy 516 CCTAAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTCTGAAGCAATTGAAAAAGCA 575  
Db 724 CCTAAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTCTGAAGCAATTGAAAAAGCA 665  
Qy 576 AATATCAGCTTTGATGATACATGTTTGAAGAGTGGGATGTAATTAATAATTCACATCTG 635  
Db 664 AATATCAGCTTTGATGATACATGTTTGAAGAGTGGGATGTAATTAATAATTCACATCTG 605  
Qy 636 ATCAATGTCTTAATGTGGGAATCTTGAAGAAAGTCAAGTGTGAGATTAACATGAATTTG 695  
Db 604 ATCAATGTCTTAATGTGGGAATCTTGAAGAAAGTCAAGTGTGAGATTAACATGAATTTG 545  
Qy 696 CTCAGCTTGGCAGCAGCAATCATTTGGGGAAGGATCTTACAGTTGCTGATGGCAGAGTG 755  
Db 544 CTCAGCTTGGCAGCAGCAATCATTTGGGGAAGGATCTTACAGTTGCTGATGGCAGAGTG 485  
Qy 756 GATGAATGAGCCAGGATATAGTTAAATACAAACATACATGAGGAATCTAGTAAACAA 815  
Db 484 GATGAATGAGCCAGGATATAGTTAAATACAAACATACATGAGGAATCTAGTAAACAA 425

QY	816	CAGCAGAGAAACATCAGTATTCAGCAGCGTGCAGCAGGAGATATTCAGCGCCAGAGC	875
Db	424	CAGCAGAGAAACATCAGTATTCAGCAGCGTGCAGCAGGAGATATTCAGCGCCAGAGC	365
QY	876	CGAGGAGAACCCCGCTCCCTGAGGAGACCTGTCAAACCTCTTCAAACACACACAGCG	935
Db	364	CGAGGAGAACCCCGCTCCCTGAGGAGACCTGTCAAACCTCTTCAAACACACACAGCG	305
QY	936	CCTGCCAGATGAGCTCGCTGCTATTCAGCGCCAGATAAACACTTACTGCGCAACATC	995
Db	304	CCTGCCAGATGAGCTCGCTGCTATTCAGCGCCAGATAAACACTTACTGCGCAACATC	245
QY	996	AAGGAGTTCATCTGCCCAAACTTAGGCAAGCTCTTCATGGCCAGGCTCTTCAAGATAC	1055
Db	244	AAGGAGTTCATCTGCCCAAACTTAGGCAAGCTCTTCATGGCCAGGCTCTTCAAGATAC	185
QY	1056	AACAACTAAGAAAGGAGTTTCCAGAAAGAAAGTTAAACATCTTTGAAGTCACACC	1115
Db	184	AACAACTAAGAAAGGAGTTTCCAGAAAGAAAGTTAAACATCTTTGAAGTCACACC	125
QY	1116	AGGCAACTCTTGGAGAAATATATTTGCATATTTGAAAGCAGAGATTTCTTTAGTG	1175
Db	124	AGGCAACTCTTGGAGAAATATATTTGCATATTTGAAAGCAGAGATTTCTTTAGTG	65
QY	1176	TCATTGGCGATTTTGGCTATACAGTGTCTTTCTAGCCATAATAAAAAA	1232
Db	64	TCATTGGCGATTTTGGCTATACAGTGTCTTTCTAGCCATAATAAAAAA	8
RESULT 11			
BM917227			
LOCUS			
DEFINITION			
5', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

ORIGIN		NIH_MGC Library."	
Query Match		76.9%; Score 984; DB 4; Length 1073;	
Best Local Similarity		97.1%; Pred. No. 1.6e-190;	
Matches 1032; Conservative		0; Mismatches 28; Indels 3; Gaps 3;	
QY	82	AAGGCAAGCGCGCTCGGAGATTACGCGTGAAGCAAGTGCAGATAGATGGCTTTGTGG	141
Db	1	AAGGCAAGCGCGCTCGGAGATTACGCGTGAAGCAAGTGCAGATAGATGGCTTTGTGG	60
QY	142	TATTAAGAATAAACAACATTAACAAGAAGAGCAAGCAAGTGAAGTTGTTCAAGGAG	201
Db	61	TATTAAGAATAAACAACATTAACAAGAAGAGCAAGCAAGTGAAGTTGTTCAAGGAG	120
QY	202	TGCTTTGGTCTGGTTGTAGAAGATCGGCTTGAATTAACCAACTGCTTTCCCTCCCTC	261
Db	121	TGCTTTGGTCTGGTTGTAGAAGATCGGCTTGAATTAACCAACTGCTTTCCCTCCCTC	180
QY	262	AGCACACAGGATGATGCTGCTTTGTAGTGAAGTCCAATATCAGATGGAATGATCGGA	321
Db	181	AGCACACAGGATGATGCTGCTTTGTAGTGAAGTCCAATATCAGATGGAATGATCGGA	240
QY	322	GCCTTCGCATGTAAACATTTGATCATCTTCACTGGGCTGGTATCAGTCCACATCTATG	381
Db	241	GCCTTCGCATGTAAACATTTGATCATCTTCACTGGGCTGGTATCAGTCCACATCTATG	300
QY	382	GCTCATCTGTTACCGGGCACTCTCGGACTCTCAGTTAGTTACCAAGCATGCCATTAAG	441
Db	301	GCTCATCTGTTACCGGGCACTCTCGGACTCTCAGTTAGTTACCAAGCATGCCATTAAG	360
QY	442	AATCTGTCTCTCTCAATTTATGATCCCAATAAAGTGGGCTGGTATCAGTCCACATTAAG	501
Db	361	AATCTGTCTCTCTCAATTTATGATCCCAATAAAGTGGGCTGGTATCAGTCCACATTAAG	420
QY	502	CATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTCGAAG	561
Db	421	CATACAGACTGACTCTCTAACTGATGGAAGTTTGTAAAGAAAGGATTTTCCCTCGAAG	480
QY	562	CATTGAAAGAAACAAATATACCTTTGATGATCATGTTTGAAGAAAGTGGCATTAATTA	621
Db	481	CATTGAAAGAAACAAATATACCTTTGATGATCATGTTTGAAGAAAGTGGCATTAATTA	540
QY	622	AAAAATTCATCTGATCAATGTCTTAATGTGGAACTTTGAAAGAAAGTGGCATTTGAG	681
Db	541	AAAAATTCATCTGATCAATGTCTTAATGTGGAACTTTGAAAGAAAGTGGCATTTGAG	600
QY	682	ATAAACATGAATGCTCAGCCTTGCAGCAAGTCAATTTGGGGAAGATCTACAGTTGC	741
Db	601	ATAAACATGAATGCTCAGCCTTGCAGCAAGTCAATTTGGGGAAGATCTACAGTTGC	660
QY	742	TGATGCAGAGTGGATGAAATGAGCCAGGATATAGTTAAATACAAACATACATGAGA	801
Db	661	TGATGCAGAGTGGATGAAATGAGCCAGGATATAGTTAAATACAAACATACATGAGA	720
QY	802	ATACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCCAGCAGGAGAATA	861
Db	721	ATACTAGTAAACACAGCAGCAGAAACATCAGTATCAGCAGCGTCCAGCAGGAGAATA	780
QY	862	TGCAGCGCCAGCGGAGGAGAACCCCGCTCCCTGAGGAGAGCTGTCTCAAACTCTTCA	921
Db	781	TGCAGCGCCAGCGGAGGAGAACCCCGCTCCCTGAGGAGAGCTGTCTCAAACTCTTCA	839
QY	922	AACCAACACAGCGGCTGCGGAGTGGCTGCTCATTGAGGCGCAGATTAACACT	980
Db	840	AACCAACACAGCGGCTGCGGAGTGGCTGCTCATTGAGGCGCAGATTAACACT	899
QY	981	TACTGCCAGAACATCAAGGAGTTTCACTGCCCAAACTTAGGCAAGCTTCTCATGG-CCCA	1039
Db	900	TACTGCCAGAACATCAAGGAGTTTCACTGCCCAAACTTAGGCAAGCTTCTCATGG-CCCA	959
QY	1040	GGCTCTTCAAGAAATACAAACACTAAGAAAGGAGTTTCCAGAAAGAGTTTAAACATGAA	1099



Db	960	GGCTCTTTTCAGATACCNACACCTTANAAAAAGGAGTGTTCAGAAAAGAGTTACCTTGAA	1019
Qy	1100	CTCTTGAAGTCACACCGGGCAACTCTTGGAGAATAATATTTT	1142
Db	1020	CTCTTGAAGTCCACCGGGCAACTCTTGGAAAAAATAATT	1062
RESULT 12			
AL570102/c			
LOCUS	AL570102	1051 bp	mRNA linear EST 05-APR-2004
DEFINITION	AL570102 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI006YN14 3-PRIME, mRNA sequence.		
ACCESSION	AL570102		
VERSION	AL570102.3	GI:46236332	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1051)		
TITLE	Li, W.B., Gruber, C., Jesses, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	On Feb 16, 2001 this sequence version replaced gi:31291529.		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4657.r		
	For more information about this cluster, see		
	http://www.genoscope.cns.fr/cdna?s=CSODI006DG07NP1&c=4657.r.		
FEATURES	Location/Qualifiers		
source	1. 1051		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSODI006YN14"		
	/tissue_type="PLACENTA COT 25-NORMALIZED"		
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
ORIGIN			
Query Match	75.8%;	Score 970.6;	DB 1; Length 1051;
Best Local Similarity	96.9%;	Pred. No. 8.5e-198;	
Matches	995;	Conservative	22; Mismatches 7; Indels 3; Gaps 3;
Qy	145	TAAAGATAATCAAAACATTATCAAGAAGGACCAAGGAACCTGAAGTTGTTCAAGGAGTGC	204
Db	1048	TTAAAWAATCAACATWATCAAG-ARRAGGACAAAGAACTG-AGKTGTTCAAGGAGGWC	991
Qy	205	TTTTGGGTCTGGTTGTAGAAGATCGGCTTGAAATATCAAACTGGTTTCCTTCCTCAGC	264
Db	990	TTTTGGGTCTGGTTGTAGAAGATCGGCTTGAAATATCAAACTGGTTTCCTTCCTTCMARC	931
Qy	265	ACACAGAGGATGATCGTCACTTTGATCAAGTCCAAATATCAGATCGGAATCATCGGGAGCC	324
Db	930	ACACAGAGGATGATCGTCACTTTGATCAAGTCCAAATATCAGATCGGAATCATCGGGAGCC	871
Qy	325	TTGCCATGTAAACATTGATCATCTTCCACGTGGGCTGGTATCAGTCCACATCTATGCGCT	384
Db	870	TTGCCATGTAAACATTGATCATCTTCCACGTGGGCTGSTATCAGTCCACATCTATGCGCT	811
Qy	385	CATTTCGTTACCCGGGCACTCCTCGACTCTCAGTTTAGTTTACCAAGCATGCGCAATGAAGAT	444
Db	810	CATTTCGTTACCCGGGCACTCCTCGACTCTCAGTTTAGTTTACCAAGCATGCGCAATGAAGAT	751





Query Match	75.5%; Score 966.2; DB 5; Length 1014;	
Best Local Similarity	97.6%; Pred. No. 6.7e-187;	
Matches 991; Conservative	0; Mismatches 23; Indels 1; Gaps 1;	
Qy	27	ACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGCGCAGCAGGAAAGGCAAGGC 86
Db	1	ACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCGCGCAGCAGGAAAGGCAAGGC 60
Qy	87	AAAGGCGGCTCGGAGATTCAGCGGTGAAGCAAGTGCAGATAGATGGCTTTGTGGTATTA 146
Db	61	AAAGGCGGCTCGGAGATTCAGCGGTGAAGCAAGTGCAGATAGATGGCTTTGTGGTATTA 120
Qy	147	AAGATAATCAAAATTCATCAAGNAGAGACAGCAAGNACTGAAGTTGTTCAAGGAGTGCTT 206
Db	121	AAGATAATCAAAATTCATCAAGNAGAGACAGCAAGNACTGAAGTTGTTCAAGGAGTGCTT 180
Qy	207	TTGGGCTGGTGTGAGAGATCGGCTTGAAGTTTACCACCTGCTTCCCTCAGCAC 266
Db	181	TTGGGCTGGTGTGAGAGATCGGCTTGAAGTTTACCACCTGCTTCCCTCAGCAC 240
Qy	267	ACAGAGGATGATGCTGACCTTTGATGAAGTCCAATATCAGATGGAATGATGCGGAGCCTT 326
Db	241	ACAGAGGATGATGCTGACCTTTGATGAAGTCCAATATCAGATGGAATGATGCGGAGCCTT 300
Qy	327	CGCCATGTAAACATTCATCATCTTCACTGGGCTGTATTCATGTCACATATATGGCTCA 386
Db	301	CGCCATGTAAACATTCATCATCTTCACTGGGCTGTATTCATGTCACATATATGGCTCA 360
Qy	387	TTGGTACCGGGCCTCTGAGCTCTGAGTTAGTTAGTACCAGCATGCCATGGAAGATCT 446
Db	361	TTGGTACCGGGCCTCTGAGCTCTGAGTTAGTTAGTACCAGCATGCCATGGAAGATCT 420
Qy	447	GTGCTTCTCATTTATGATCCCATAAAACTGCCAAGGATCTCTCTCACTAAAGGCATAC 506
Db	421	GTGCTTCTCATTTATGATCCCATAAAACTGCCAAGGATCTCTCTCACTAAAGGCATAC 480
Qy	507	AGACTGACTCTAAACATGATGAAGTTTGTAAAGAAAGGATTTTCCCTGGAAGCATTG 566
Db	481	AGACTGACTCTAAACATGATGAAGTTTGTAAAGAAAGGATTTTCCCTGGAAGCATTG 540
Qy	567	AAAAAGCAAAATATCACTTTGAGTACATGTTTGAAGAGTGCCTGTTGTAATTAATAAT 626
Db	541	AAAAAGCAAAATATCACTTTGAGTACATGTTTGAAGAGTGCCTGTTGTAATTAATAAT 600
Qy	627	TCACATCTGATCAATGCTCTTAATGTGGGAACCTTGAAAAGAGTCAAGCTGTTGCAGATAA 686
Db	601	TCACATCTGATCAATGCTCTTAATGTGGGAACCTTGAAAAGAGTCAAGCTGTTGCAGATAA 660
Qy	687	CATGAATGCTCAGCTTCCAGCAGCAATATCAATTTGGGGAAGATCTACAGTTGCTGATG 746
Db	661	CATGAATGCTCAGCTTCCAGCAGCAATATCAATTTGGGGAAGATCTACAGTTGCTGATG 720
Qy	747	GACAGAGTGAATAATGAGCCAGATATAGTTAAATACACACATATAGAGGAATACT 806
Db	721	GACAGAGTGAATAATGAGCCAGATATAGTTAAATACACACATATAGAGGAATACT 780
Qy	807	AGTAAACAAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCGCAGCAGAGATATGCG 866
Db	781	AGTAAACAAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCGCAGCAGAGATATGCG 840
Qy	867	CGCCAGCGCGGAGAACCCCGCTCCTCGAGGAGGACCTGTCCAAACTCTTCAAAACA 926
Db	841	CGCCAGCGCGGAGAACCCCGCTCCTCGAGGAGGACCTGTCCAAACTCTTCAAAACA 899
Qy	927	CACAGCGGCTGCCAGGATGAGCTGCTCATTTGCGGCGCAGATAAACTACTATGCG 986
Db	900	ACACAGCGGCTTGCAGGATGAATCGCTGCTCATTTGCAAGCCAGATAAACTACTATGCG 959
Qy	987	CAGAAATCAAGAGTTCACTGCCCAAACTTATAGGCAAGCTCTTCACTGGCCGAGG 1041
Db	960	CGGAACTCAAGAGTTCACTGGCCCAAACTTATAGGCAAGCTCTTCACTGGCCGAGG 1014

[illegible]

Search completed: January 5, 2005, 16:08:28  
Job time : 4413 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:34:10 ; Search time 80.6851 Seconds  
(without alignments)  
289.321 Million cell updates/sec

Title: US-10-017-327-2

Perfect score: 1812

Sequence: 1 MASRKEGTGSTATSSSTAG.....FTAQLGKLFMAQALQEYNN 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	42.5	346	US-09-270-767-45452	Sequence 45452, A
2	334	18.4	311	US-09-248-796A-16692	Sequence 16692, A
3	158.5	8.7	357	US-09-149-476-755	Sequence 755, App
4	156.5	8.6	261	US-09-248-796A-19618	Sequence 19618, A
5	149.5	8.3	342	US-09-270-767-45174	Sequence 45174, A
6	124	6.8	455	US-09-538-092-123	Sequence 123, App
7	114	6.3	342	US-09-248-796A-19454	Sequence 19454, A
8	108	6.0	707	US-09-919-039-278	Sequence 278, App
9	108	6.0	707	US-09-538-092-993	Sequence 993, App
10	108	6.0	1201	US-09-328-352-5990	Sequence 5890, Ap
11	107.5	5.9	305	US-09-248-796A-20913	Sequence 20913, A
12	105.5	5.8	195	US-09-213-293D-19	Sequence 19, Appl
13	105.5	5.8	1082	US-08-106-493A-2	Sequence 2, Appl
14	105.5	5.8	1082	US-08-429-264-2	Sequence 2, Appl
15	105.5	5.8	1139	US-08-832-883-2	Sequence 2, Appl
16	105.5	5.8	1139	US-08-832-877-2	Sequence 2, Appl
17	104	5.7	1255	US-09-919-172-29	Sequence 29, Appl
18	104	5.7	2704	US-09-538-092-1260	Sequence 1260, Ap
19	102.5	5.7	383	US-09-248-796A-16746	Sequence 16746, A
20	101.5	5.6	471	US-08-866-928B-1	Sequence 1, Appl
21	101.5	5.6	471	US-09-685-836-1	Sequence 1, Appl
22	101.5	5.6	738	US-09-538-092-818	Sequence 818, App
23	101	5.6	321	US-09-270-767-33569	Sequence 33569, A
24	101	5.6	321	US-09-270-767-48786	Sequence 48786, A
25	100.5	5.5	297	US-09-248-796A-22393	Sequence 22393, A
26	100.5	5.5	353	US-09-248-796A-17342	Sequence 17342, A
27	100.5	5.5	712	US-08-468-576B-17	Sequence 17, Appl

28	100.5	5.5	712	2	US-08-468-579B-17	Sequence 17, Appl
29	100.5	5.5	712	3	US-08-468-577B-17	Sequence 17, Appl
30	99.5	5.5	1279	3	US-09-724-517-2	Sequence 2, Appl
31	99.5	5.5	1279	4	US-09-641-807A-2	Sequence 2, Appl
32	99.5	5.5	1279	4	US-09-723-036-2	Sequence 2, Appl
33	99	5.5	563	3	US-09-134-001C-3172	Sequence 3172, Ap
34	98.5	5.4	286	4	US-09-248-796A-16992	Sequence 16992, A
35	98.5	5.4	327	3	US-09-008-271A-10	Sequence 10, Appl
36	97	5.4	341	4	US-09-710-279-226	Sequence 226, App
37	97	5.4	532	4	US-09-710-279-546	Sequence 546, App
38	96	5.3	48	4	US-09-270-767-60958	Sequence 60958, A
39	96	5.3	273	4	US-09-538-092-149	Sequence 149, App
40	96	5.3	484	4	US-09-248-796A-24175	Sequence 24175, A
41	95.5	5.3	358	4	US-09-248-796A-22578	Sequence 22578, A
42	95.5	5.3	513	3	US-09-041-886-28	Sequence 28, Appl
43	95.5	5.3	530	3	US-09-041-886-29	Sequence 29, Appl
44	95.5	5.3	552	3	US-09-041-886-30	Sequence 30, Appl
45	95.5	5.3	589	3	US-09-041-886-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1  
US-09-270-767-45452  
; Sequence 45452, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45452  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-45452

Query Match 42.5%, Score 771, DB 4, Length 346;  
Best Local Similarity 45.2%, Pred. No. 5e+65;  
Matches 149, Conservative 70, Mismatches 109, Indels 2, Gaps 1;

Qy	19	AGAAGKGGKGGSDSAVKQVIGLVLTIIKHVQEGGCTEVVQGVLLGLWVEDRL	78
Db	10	ANRANRHAARTSDSDNTINYQCDGLAVMRKVKHCHESSNMDLAQGLLVVDKCL	69
Qy	79	TNCFPPFOHTDDADDFEVOYQEMEMRSLRVNIDHLHVGVYQSTYYGSFVTRALLDSQ	138
Db	70	TNCFPPFK--SGDETDEMYQLTVMRRLRVNVDHLHVGVYQSDVGNLSMALLESQY	127
Qy	139	SYQHAIBESVLLIYDPIKTAGGSLKAYRLTPKLMVECKEKDPSPEALKKANITFEYMF	198
Db	128	HYQTSIBESVVVVYDTQKSSRGFLCLKAYRLTPQAIQYKGDGFTPEAFRLTKVGYENL	187
Qy	199	EEVPVINKSHLINVLMWELEKSAVADKHELLSLASSNHLGKNIQLLMDRVDMSQDIV	258
Db	188	AEIPVINKSPITNIMSELSNELLPEKGNHFLDGTATVLENQMRSLIERVDELYQAV	247
Qy	259	KNTYMTNRSKQOQKHQYQORQENNRQSRGEPPLPEEDLSKLPKPPPPARMDSLL	318
Db	248	RYNKYQVVFQDTEKHALAKLAENAVRTSKGPTVPEEVIKQFPFMTAPNRLTATI	307
Qy	319	IAGQINTYQCNKEFTAQNLGKLFMAQALQ	348
Db	308	TSQGINTHAQIAQFCQSLAKLFITESLQ	337

RESULT 2  
US-09-248-796A-16692  
; Sequence 16692, Application US/09248796A

```
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 16692
/ LENGTH: 311
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ US-09-248-796A-16692

Query Match      18.4%; Score 334; DB 4; Length 311;
Best Local Similarity 26.7%; Pred. No. 1.5e-23;
Matches 79; Conservative 76; Mismatches 133; Indels 8; Gaps 5;

QY      34 SAVKQVQIDGLVLLKIKHYQEGGTEVQGVLLGLVDRLEITNCFFPPQHTD-DA 92
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
13 ATLNTVHISSSVLLSLRHTSE--HYPQLFSGALIGFEDESILDVTHGFPYPDPQYEGG 70
QY      93 DFDE---VQYQEMMRSLRHVNIDHLHVQWYQSTVYGSFVTRALLDSQFSYQHALESVV 149
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
71 SFRSRIQGHYQKDLLENFKLGYGIEFLGWFQSTVSGNFVNQIIEG-LAQQQIINSAF 129
QY      150 LIYDPIKTAQSLSKAYRLPKLMEVCKEDFSPALKKANIITFEYMFEEVPIVKNSH 209
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
130 ILINLSSVGQEVSIKALRSTGFNNAVYDGVKWKSKDLESNKISYLNIFEELNLEISNQK 189
QY      210 LINVLWLELKKSAVADRHELLSLASS-NHLGKNLQLLMDRVDSMDQIVKYNTRYMNTS 268
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
190 LVDVTLSSLSLKPSTESLDVLNLSNQATGQLESLSGQVDSFNVDQNNFNRYQRYQ 249
QY      269 KQQQKHQYQRRQOENNRQSRGEPPLPEEDLSKLFKPPOPPARMDSLIAGQIN 324
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
250 KEQSKIQQKQKQKLENLERAKKEKELDIETEWKTIKLFNFSRYNNMLYSKAID 305

RESULT 3
US-09-149-476-755
/ Sequence 755, Application US/09149476
/ Patent No. 6420526
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 186 Human Secreted proteins
/ FILE REFERENCE: P2002P1
/ CURRENT APPLICATION NUMBER: US/09/149,476
/ CURRENT FILING DATE: 1998-09-08
/ EARLIER APPLICATION NUMBER: PCT/US98/04493
/ EARLIER FILING DATE: 1998-03-06
/ EARLIER APPLICATION NUMBER: 60/040,162
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,333
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/038,621
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,626
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,334
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,336
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,163
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/047,600
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,615
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,597
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,502
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,633
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,583
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,617
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,618
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,503
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,592
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,581
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,584
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,500
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,587
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,492
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,598
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,613
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,582
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,596
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,612
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,632
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,601
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,580
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,568
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,314
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,569
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,311
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,671
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,674
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,669
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,312
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,313
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,672
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,315
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/048,974
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/056,886
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,877
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,889
/ EARLIER FILING DATE: 1997-08-22
```



```
; SEQ ID NO 19618
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (28),(126)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-19618

Query Match      8.6%; Score 156.5; DB 4; Length 261;
Best Local Similarity 28.2%; Pred. No. 8.9e-07;
Matches 46; Conservative 34; Mismatches 58; Indels 25; Gaps 7;

QY 23 GKGGKGGSDSA-----VKQVIDGLVLLKIKHYQEEGGQTEVQVGLGLVVED 74
Db 32 GPGAGLGGAAPTQSDPAIDNSETVYISLLALKXKH-GRAGVPMVM-GFMLGEFVDD 89

QY 75 -RLBITNCFPPFQ-----HTEDDADFDEVQYQMEMMRSRLRHVNIDHLHVGWYQS-TYY 125
Db 90 FTIHVDVFAFPQSGTGVSEAVDDV-----FQTKMDMLRXTGRDQMVVGVHSHPGF 143

QY 126 GSFVTRALLDSQFSYQHAIBESVLLIYDPIKTAQGSLSLKAYR 168
Db 144 GCMLSSVDVNTQSFQELNKRVAWIDPIQSVKGVVIDAFR 186

RESULT 5
US-09-270-767-45174
; Sequence 45174, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45174
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45174

Query Match      8.3%; Score 149.5; DB 4; Length 342;
Best Local Similarity 21.3%; Pred. No. 6.1e-06;
Matches 57; Conservative 65; Mismatches 114; Indels 31; Gaps 11;

QY 37 KOVIDGLVLLKIKHYQEEGGQTEVQVGLGLVVEDRLITNCFPPF---QHTEDDAD 93
Db 74 KDIKISALLKVMH--ARSGGTLEVMGLMKGVEDNTMIVMDAFALPVEGTETRVNAQ 131

QY 94 FDEVQIQMEMMRSRLRHVNIDHLHVGWYQS-TYYGSFVTRALLDSQFSYQHAIBESVLLIY 152
Db 132 AQAYEYMTAYMEAAKEVGRMEHAGVGHVSHPGYGCWLSGIDVSTQMLNQTQEPFVAIVV 191

QY 153 DPIKT-AQGSLSLKAVRLTPKLMVECKE--DFSPKALKKANITFEYMPFEEPIVKNKH 209
Db 192 DPTVTSAGKVCIGAFRTYKPKPNEBSEYQTPLNK-----IEDFGVHCKQYY 243

QY 210 LINVLME--LEKK--SAVADKHELLSLASSNHLGNLQLLMDRVDMSQDIVKNTVM- 264
Db 244 PLEISYFKSALDRLLDSLWNKYVWNTLGSSGLL-TNTEVTTQIIMDLSEKLEQSFNFIG 302

QY 265 RNTSKQQQKHQYQRRQENMQRSR 291
Db 303 RGT-----DVNKRSEDKLSKATR 321

RESULT 6
US-09-538-092-123
; Sequence 123, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurafatSeqformatter Version 0.9
; SEQ ID NO 123
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YDL216C
US-09-538-092-123

Query Match      6.8%; Score 124; DB 4; Length 455;
Best Local Similarity 21.4%; Pred. No. 0.0024;
Matches 58; Conservative 46; Mismatches 105; Indels 62; Gaps 12;

QY 50 IKHYQEEGGQTEVQVGLGLVVEDRLITNCFPPF---QHTEDDADFDEVQYQMEMMRS 106
Db 96 ITHYAVRGGNIEIM-GILMGFTLKDNIVVMDCFNLPVVGTETRVNAQLESYEMVQY--- 151

QY 107 LRHVNDIHLH-----VGVYQS-TYYGSFVTRALLDSQFSYQHAIBES 147
Db 152 -----IDEMYNHNDGGDRDYKAKLVNVGVFPHSHPGYDCWLSNIDIDITQDLNQRFPDPY 206

QY 148 VVLIYDPIKTAQGS-LSLKAVRLTPKLMVECKEKFSPKALKKANITFEYMPFEEPIVKNKH 206
Db 207 VAIVVDPLKSLDKILRMGAFRTIES-----KSDNNSATS-----YYELETITIF 250

QY 207 NSHLINVLMELEKKSAV-----ADKHELLSLASSNHLGNLQLLMDRVDMSQDIVKYN 261
Db 251 DSELNRL---PETKLNLCVIEDDESEISLNRLIDSMKQVSYLMDSKNVTR--IKLA 305

QY 262 TYMRNTSKQQQKHQYQRRQEE-----NMQR 288
Db 306 TTSERVSNNKKNIDYQNRSTRSQFCLNTQR 336

RESULT 7
US-09-248-796A-19454
; Sequence 19454, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19454
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19454

Query Match      6.3%; Score 114; DB 4; Length 342;
Best Local Similarity 21.3%; Pred. No. 0.014;
```



Matches	54;	Conservative	45;	Mismatches	101;	Indels	54;	Gaps	9;
Qy	10	STATSSSTAGAAAGKGGKGGSDSAVKVQIQDGLVWLKIIKHYYQEEGQGTGV-VQGVL	68						
		:							
		:							
		:							
		:							
Db	15	TTATSTNELA-----LLDKSVVVSPLVLVVDHYNRAKAKSKRVAWGVL	60						
		:							
		:							
		:							
		:							
Qy	69	GLVVEDRLKLTCTPFPHQHTEDDAD---FDEVOYQMEMMRLSHRVNIDHLHVGWGYSTY	124						
		:							
		:							
		:							
Db	61	GDNSTDIKVTNSYAIP- FEDEKQPGVWFJLHNFIDSMGFKNKINAKKELIGWHS--	117						
		:							
		:							
		:							
Qy	125	YGSFVTRALLDSQFSYQHAIEESVVLVYDPIKTAGSLSLKAYRLTKPLMEVCKEKFDP	184						
		:							
		:							
		:							
Db	118	-GPKLKPSDLKINEVFRRYTDNPLLIVD-----VQPREVGIPITDAYFAV	161						
		:							
		:							
		:							
Qy	185	EALKKANITTEYMEFEVPIVTKNSHLINVLWLEWEKKSAVADKHELLSLASSNLGNLQ	244						
		:							
		:							
		:							
Db	162	DDIKGDSAAEKTTHVPSLI-----EAEAEIEGVHEHLRDI--RDQAGNLS	208						
		:							
		:							
		:							
Qy	245	LLMDRVDMSQDIV	258						
		:							
		:							
		:							
Db	209	L---RVSETHOSLL	219						

```

RESULT 8
US-09-919-039-278
; Sequence 278, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 278
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6727066 2457215CD1
US-09-919-039-278

```

Query Match	6.0%; Score 108; DB 4; Length 707;
Best Local Similarity	21.1%; Pred. No. 0.15;
Matches	46; Conservative 32; Mismatches 66; Indels 74; Gaps 8;
Qy	128 FVTRALDLSFYSYHAJEEFVLLYDPIK-TAQSLSLKAYRLTPKLMVECKE-----179
Db	381 YVSNELLEAFSQGPPIERAVIVDDRGRTGKGI VEFASKPAARKAFERCSGVFLTT 440
Qy	180 -----KDFSPALKKANI-----TFPYMFEFPIVIKNS 208
Db	441 TPRPVIVEPLEQLDDEGLPEKLAQKNPMYQKERETTPRFAQHGTFEYYSQ-----492
Qy	209 HLINVLWV----ELEKKSAYADRHLLSLASSNHLGNLQLMDRVDEMSDIVKYNTYM 264
Db	493 -----RWKSLDENEKO-----QREYVKNMKADKLESEMED--AYHEHQ 531
Qy	265 RNTSKQQQKHQYQRRQOE--NMQRQSRGEPPLPEED 300
Db	532 ANLLRQDLMRQEBELRRMEELHNOEMOKRKEMOLRQEB 569

RESULT 9  
US-09-538-092-993  
; Sequence 993, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic

```

; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 993
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P23246
; US-09-538-092-993

```

Query Match	6.08; Score 108; DB 4; Length 707;
Best Local Similarity	21.18; Pred. No. 0.15;
Matches	46; Conservative 32; Mismatches 66; Indels 74; Gaps 8;
Qy	128 FVTRALDSDFSYQHATIEESWLIYDPIK-TAQSLSLKAYRLTPKLMVEVCKE-----179
Db	381 YVSNELLEAFSGFGPIERAVVVDGRSTGKIVFEFASKPAARKAFERCSGCVFLLT 440
Qy	180 -----KDFSPALKKANI-----TFPYMFEFVPIVKN 208
Db	441 TPRPVIVEPQLDDEDLGPEKLAQKNPMYQKERETPPRFAHQHGTFFEYSQ-----492
Qy	209 HLINVLMW----BLEKSAVADKHLLSLASSNHLGKNLQLLMDRVDEMSDIVKYNTYM 264
Db	493 -----RWKSLDEMKEQ-----QREQVEKNMKDARDKLESEMED--AYHEHQ 531
Qy	265 RNTSKQQQKHQYQRRQOE--NQQRSGRGEPPPLEED 300
Db	532 ANLLRQDLMRQBELRRMEELHMQEMQKRKEMQLRQEE 569

```

RESULT 10
US-09-328-352-5890
; Sequence 5890, Application US/09328352
; Patent NO. 6562958
; GENERAL INFORMATION:
; APPLICANT: GARY L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5890
; LENGTH: 1201
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5890

```

	Query Match	6.0%;	Score 108;	DB 4;	Length 1201;
	Best Local Similarity	22.1%;	Pred. No. 0.32;		
	Matches	76;	Conservative	58;	Mismatches 100; Indels 110; Gaps 18;
Qy	51	KHQBEGQTEVVQGVLLGLHVEDRL	LEITNCFFPPQHTEDD-----	ADFEVQYQMEMMR	105
Db	539	KHTQASAQ-----	EDQLK-TVC---QLAEQDYQTA	TEREKLIHQILQQOR	579
Qy	106	SLRHVNIDHLHVGWYQ-----	STVY-----GSFVTRLLDSQFSYQ	H---AIBESV	149
Db	580	LLHTENIEQRANLKGEAC	LCVCGSTHHPPYRIDDSAVSKALFD	LQOQQEHQVALEQTKF	639
Qy	150	LIYDPKTAQGSLSLKAYRLTP	KLMVECKEKFSP	EALKKANITTFYMF	FEVPVIVKNSH 209

Db 640 NVW---QTCQALH---QCRALFQVQK-----YLAQ---LQVQTH 672  
QY 210 LINVLMWELEKKSADVADKHELLS-----LASSNHLGKNLQLLMDRVD-----EMSDIVK 259  
Db 673 ----LQBLEQAFNLNQLHI ELNQAPEQILOTLNELRQATQTALSFDSENRLRTOAIKQ 728  
QY 260 VNTYMR-----NTSKQOQO-----KHQOQROQOENMOR-----QSRGE 293  
Db 729 HNLQVQTTORNESLLNTAQOQOQVQVHVECLSETEQAWQOQASSQTAKQTWAILDARAK 789  
QY 294 PPLPEDLSKLFKPPPPARMDSLLI---AGQINTYCNIKEFT 334  
Db 789 QLELQELQSRFEQOQOQELKMSLSLEQMTQIDQIDQLQEI 832

RESULT 11  
US-09-248-796A-20913  
; Sequence 20913, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 20913  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-20913

Query Match 5.9%; Score 107.5; DB 4; Length 305;  
Best Local Similarity 19.1%; Pred. No. 0.051;  
Matches 54; Conservative 47; Mismatches 93; Indels 89; Gaps 11;  
QY 97 VOYQMEMRSLRHVDHILHGVYQS--TYYGSVFTRALLDSQFSYQHAIEESVVLIDPT 155  
Db 18 VEYLTERQQLNGKNEENI-VGYTHSPGYGCLSGIDVSTQSLNGQFQDPYLAIVVDV 76  
QY 156 KT-AQGSLSLKAVRLTP-----KLMEVCKEKFSP 185  
Db 77 KTLQKQKEVIGAFRTYPEGSOQPSMTNTRKQNKPHNSGANANKILPKSKQKDFGSH 136  
QY 186 ALKANITTFYMF-----ERVIVIKNSHLINVLWLEKKSADV-----KHSL 230  
Db 137 ADKYSLDIEIFTSSWDKVIEMLKDB---DSLTV--MKNLIVDSNNNDKILGIRKDEI 190  
QY 231 LS-----LASSNHLGKNLQLLMDRDEM-----SQDIVKY 260  
Db 191 RSLIELKNVLIQGNHNADEGETIFDLBQLKIQANTPKFMDLKLTTMKFDSNFESVLY 250  
QY 261 NTYWRNTSKQOQKHQVQORROQENMORSGRPPPLPEEDLSK 303  
Db 251 KRLKKTKQSTTK-----KNRKLSTDDIDETMLDESLEK 287

RESULT 12  
US-09-213-293D-19  
; Sequence 19, Application US/09213293D  
; Patent No. 6384299  
; GENERAL INFORMATION:  
; APPLICANT: GUTIERREZ-ARMENTA, CRISANTO  
; APPLICANT: SANZ-BURGOS, ANDRES P.  
; APPLICANT: XIE, QI  
; APPLICANT: LOPEZ, PAULA S.  
; TITLE OF INVENTION: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS

; FILE REFERENCE: 604-469  
; CURRENT APPLICATION NUMBER: US/09/213,293D  
; CURRENT FILING DATE: 1998-12-14  
; PRIOR APPLICATION NUMBER: PCT/ES96/00130  
; PRIOR FILING DATE: 1996-06-13  
; PRIOR APPLICATION NUMBER: PCT/EP97/03070  
; PRIOR FILING DATE: 1997-06-12  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-213-293D-19

Query Match 5.8%; Score 105.5; DB 3; Length 195;  
Best Local Similarity 24.9%; Pred. No. 0.041;  
Matches 52; Conservative 24; Mismatches 86; Indels 47; Gaps 9;  
QY 154 PIKTAQGSLSL---KAYRLTP-KLMEVCKEKFSPKALKANITFEYMFEEVPIVKNSH 209  
Db 3 PKRTS--SLSLFRKYVYHLAAVRLDLCALKDLSDELKRWTCFEFSIIQCPELMMDRH 60  
QY 210 LINVLMWELEKKSADVADKHELLSLASSNHLGKNLQLLM-----DRVDEMSQDIVKYN 261  
Db 61 LQQLLCAIYVNAKYKE-----DKSFQNIWRCTYQOARSQVYRSVLIKKG 108  
QY 262 TYWRNTSKQOQKHQVQORROQENMORSGRPPPLPEEDLSKLFKPPQP-----PARMD 315  
Db 109 RKRNRSGSDSRSH--QNSPTLNKDRTSRDSQSPWRS--SSTLPVPQPPSSAAPTPTRLT 164  
QY 316 SL-----LIAGQINTYCNIKEF 333  
Db 165 GANSMDMEERGGDLIQFYNNIYIKQIKTP 193

RESULT 13  
US-08-106-493A-2  
; Sequence 2, Application US/08106493A  
; Patent No. 5457049  
; GENERAL INFORMATION:  
; APPLICANT: Antonio Giordano  
; TITLE OF INVENTION: "TUMOR SUPPRESSOR PROTEIN PRE2,  
; TITLE OF INVENTION: RELATED GENE PRODUCTS, AND DNA ENCODING  
; TITLE OF INVENTION: THEREFOR"  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Temple University-Of The Commonwealth  
; ADDRESSEE: System of Higher Education  
; STREET: 406 University Services Building  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/106,493A  
; FILING DATE: August 12, 1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mullins, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 6056-188  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549

```
; TELEX: No. 5457049e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1082 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-106-493A-2

Query Match          5.8%; Score 105.5; DB 1; Length 1082;
Best Local Similarity 24.9%; Pred. No. 0.48;
Matches 52; Conservative 24; Mismatches 86; Indels 47; Gaps 9;

Qy 154 PIKTAQGSLSL---KAYRLTP-KLMEVCKEKDFSEALKKANITFEYMFEEVPIVKNSH 209
Db 773 PRKTS--SLSLFFRKVYHLAAVRLDLCALDISDELKTKIWTCTFEFSIIQCPELMMDRH 830

Qy 210 LINVLMELEKKSAAVADKHELLSLASSNHLGKNLQLLM-----DRVDEMSQDIVKYN 261
Db 831 LDQLLMCAIYMAKVTK-----DKSFQNMRCYRTQPARSQVYRSVLKKG 878

Qy 262 TYMNTSKQQQKHQYQORQENMQRSRGEPLPEEDLSKLPKPOP-----PARMD 315
Db 879 RKRNSGSSDSRSH--QNSPTLNKDRTSRDSPPWRS--SSTLPVPQSPAPTPTRLT 934

Qy 316 SL-----LIAGQINTYCONIKEF 333
Db 935 GANSDBEERGDLIQFYNNIYKIQITF 963

RESULT 14
US-08-429-264-2
; Sequence 2, Application US/08429264
; Patent No. 5532340
; GENERAL INFORMATION:
; APPLICANT: Antonio Giordano
; TITLE OF INVENTION: "TUMOR SUPPRESSOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,264
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/106,493
; FILING DATE: August 12, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, D.A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-188 DII
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5532340e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1082 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-429-264-2

Query Match          5.8%; Score 105.5; DB 1; Length 1139;
Best Local Similarity 24.9%; Pred. No. 0.51;
Matches 52; Conservative 24; Mismatches 86; Indels 47; Gaps 9;

Qy 154 PIKTAQGSLSL---KAYRLTP-KLMEVCKEKDFSEALKKANITFEYMFEEVPIVKNSH 209
Db 830 PRKTS--SLSLFFRKVYHLAAVRLDLCALDISDELKTKIWTCTFEFSIIQCPELMMDRH 887

Qy 210 LINVLMELEKKSAAVADKHELLSLASSNHLGKNLQLLM-----DRVDEMSQDIVKYN 261
Db 888 LDQLLMCAIYMAKVTK-----DKSFQNMRCYRTQPARSQVYRSVLKKG 935
```

[illegible]

Search completed: December 30, 2004, 20:53:54  
Job time : 82.6851 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:25:59 ; Search time 410.519 Seconds  
(without alignments)  
307.593 Million cell updates/sec

Title: US-10-017-327-2

Perfect score: 1812

Sequence: 1 MASRKEGTGSTATSSSTAG.....FTAQLGKLFMAQALQEQYNN 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_23Sep04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1812	100.0	352	5	ABB08360 Human eIF
2	1812	100.0	352	7	ABR82212 Human ant
3	1808	99.8	352	7	ADJ68541 Human hea
4	1794	99.0	368	7	ADC31312 Human nov
5	1792	98.9	352	5	ABB08366 Human can
6	1789	98.7	347	7	ADE29059 Human EIF
7	1785	98.5	352	5	ABB08367 Human can
8	1773	97.8	352	5	ABB08368 Human can
9	1765	97.4	352	5	ABB08369 Human can
10	1552	85.7	302	7	ADE29060 Human EIF
11	1275	70.4	256	3	AA56775 Human pro
12	1026	56.6	202	5	ABP42329 Human ova
13	723	39.9	307	4	ABB64255 Drosophil
14	627.5	34.6	334	3	AA616828 Arabidops
15	627.5	34.6	334	3	AA616828 Arabidops
16	627.5	34.6	337	3	AA616827 Arabidops
17	627.5	34.6	337	3	AA616827 Arabidops
18	478	26.4	250	3	AA616829 Arabidops
19	478	26.4	250	3	AA616829 Arabidops
20	304	16.8	59	5	ABP41576 Human ova
21	183	10.1	310	8	ADO19863 Human PRO
22	183	10.1	310	8	ADO19863 Human PRO
23	183	10.1	315	3	AA656660 Human pro
24	180.5	10.0	306	6	ABR53088 Protein s
25	180.5	10.0	306	7	ADK62986 Disease t

26	177.5	9.8	308	4	ABB66760 Drosophil
27	176.5	9.7	312	5	ABP73468 Candida a
28	167.5	9.2	332	7	ADB70086 C. neofo
29	166.5	9.2	307	2	AA608442 Z. mays p
30	166.5	9.2	308	2	AA608454 Yeast Rad
31	164.5	9.1	308	7	ABM73834 DNA clone
32	164.5	9.1	309	2	AA608443 Soybean p
33	158.5	8.7	357	5	ABG95615 Human nov
34	158.5	8.7	357	6	ABO34809 Fragment
35	158.5	8.7	357	7	ADI23470 Novel hum
36	158.5	8.7	357	8	ADH74472 Human sec
37	156.5	8.6	308	3	AA651138 Arabidops
38	156.5	8.6	308	8	ADI61411 A. thalia
39	156.5	8.6	308	8	ADO02355 Thalecres
40	156.5	8.6	375	2	AA608453 Soybean J
41	155	8.6	314	3	AA624656 Arabidops
42	154.5	8.5	305	6	ABJ25834 Aspergill
43	154.5	8.5	308	3	AA615852 Arabidops
44	153.5	8.5	287	3	AA654466 Zea mays
45	151.5	8.4	335	6	ABJ26434 Aspergill

ALIGNMENTS

RESULT 1					
ABB08360					
ID	ABB08360	standard; protein; 352 AA.			
XX	AC	ABB08360;			
XX	XX				
DT	07-MAY-2002	(first entry)			
DE	Human eIF3 amino acid sequence.				
XX	Human; melanoma antigen eukaryotic initiation factor 3; eIF3;				
KW	ovarian cancer; MHC; cytosstatic; immunomodulator; immune effector cell;				
KW	anti-cancer; vaccine.				
XX	XX				
OS	Homo sapiens.				
XX	XX				
FH	Key	Location/Qualifiers			
FT	Region	242..250			
FT		/note= "natural epitope of human cancer antigen eIF3"			
FT	Binding-site	242			
FT		/note= "HLA-2 binding residue"			
FT	Binding-site	243			
FT		/note= "HLA-2 binding residue"			
FT	Binding-site	244..249			
FT		/note= "T-cell receptor (TCR) binding domain"			
FT	Binding-site	250			
FT		/note= "HLA-2 binding residue"			
WO200192307-A2.					
06-DEC-2001.					
30-MAY-2001; 2001WO-US017456.					
31-MAY-2000; 2000US-0209391P.					
17-AUG-2000; 2000US-0226258P.					
20-DEC-2000; 2000US-0257008P.					
(GENZ ) GENZYME CORP.					
Nicolette CA;					
WPI; 2002-139606/18.					
N-PSDB; ABA97211.					
New therapeutic compounds useful against human ovarian cancer, for					
modulating immune response in a subject, and for generating antibodies					
that specifically recognize and bind to these molecules.					

XX PS Disclosure; Page 63-64; 68pp; English.

XX CC The invention relates to novel therapeutic compounds, that are designed

XX CC to enhance binding to MHC molecules and to enhance immunoregulatory

XX CC properties relative to their natural counterparts. The activity of the

XX CC compounds of the invention may be described as cytostatic and

XX CC immunomodulatory. The compounds are useful against human ovarian cancer,

XX CC for modulating immune response in a subject, and for generating

XX CC antibodies that specifically recognize and bind to these molecules.

XX CC Compositions comprising the compounds are useful as components of anti-

XX CC cancer vaccines and to expand immune effector cells that are specific for

XX CC cells characterized by expression of antigen eIF3 (melanoma antigen

XX CC eukaryotic initiation factor). The peptides or polypeptides conjugated to

XX CC a detectable agent may be used in diagnostic procedures, such as in the

XX CC detection and purification of antibodies, and as immunogens for

XX CC production of antibodies. The polynucleotides can be used as primers for

XX CC detecting genes or gene transcripts expressed in APC to confirm

XX CC transduction of the polynucleotides into host cells. The current sequence

XX CC represents the human melanoma antigen eukaryotic initiation factor 3

XX CC (eIF3) amino acid sequence

XX SQ Sequence 352 AA;

Query Match 100.0%; Score 1812; DB 5; Length 352;

Best Local Similarity 100.0%; Pred. No. 3.4e-159;

Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRKEGTGTSATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHVQEGQGT 60

DB 1 MASRKEGTGTSATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHVQEGQGT 60

QY 61 EVVQVLLGLVVEDRLITNCFFPPQHTEDDADFDEVQYQEMMRSLRHVNIDHLHVGWY 120

DB 61 EVVQVLLGLVVEDRLITNCFFPPQHTEDDADFDEVQYQEMMRSLRHVNIDHLHVGWY 120

QY 121 QSTYGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEK 180

DB 121 QSTYGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEK 180

QY 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSAAVADKHELLSLASSNHLG 240

DB 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSAAVADKHELLSLASSNHLG 240

QY 241 KNQLQLMDRVDEMSSQDIVKNTYMRNTSKQQQKHQYQRRQENMQRSRGPPLPEED 300

DB 241 KNQLQLMDRVDEMSSQDIVKNTYMRNTSKQQQKHQYQRRQENMQRSRGPPLPEED 300

QY 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

DB 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

RESULT 2

ABR82212

ID ABR82212 standard; protein; 352 AA.

XX AC ABR82212;

XX DT 13-OCT-2003 (first entry)

XX XX Human antigen eukaryotic translation initiation factor 3 (eIF3).

XX XX Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;

XX KW cytostatic; gene therapy; human; antigen.

XX XX Homo sapiens.

XX OS

XX PN WO2003050543-A1.

XX PD 19-JUN-2003.

XX XX

XX PF 05-DEC-2001; 2001WO-US047997.

XX PR 05-DEC-2001; 2001WO-US047997.

XX PA (GENZ ) GENZYME CORP.

XX PI Nicolette CA;

XX XX WPI; 2003-532936/50.

DR N-PSDB; ACC85029.

XX Aiding in the diagnosis of a neoplastic condition, useful for treating

PT cancer and related malignancies comprises determining the amount of

PT expression of an eIF3 protein in a test sample isolated from the cell or

PT tissue.

XX Claim 6; Page 71-72; 77pp; English.

PS The invention relates to aiding in the diagnosis of a neoplastic

XX condition or susceptibility to a neoplastic condition of an animal cell

CC or tissue. The method involves determining the amount of expression of an

CC eukaryotic translation initiation factor 3 (eIF3) protein in a test

CC sample isolated from the cell or tissue, and diagnosing a neoplastic

CC condition or susceptibility to a neoplastic condition based on the amount

CC of expression of the eIF3 protein. The methods, compounds and kits are

CC useful in therapeutics, diagnostic and screening methods for human cancer

CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,

CC pancreatic or gastrointestinal cancer, or melanoma. The present sequence

CC represents the human antigen eIF3

XX SQ Sequence 352 AA;

Query Match 100.0%; Score 1812; DB 7; Length 352;

Best Local Similarity 100.0%; Pred. No. 3.4e-159;

Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRKEGTGTSATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHVQEGQGT 60

DB 1 MASRKEGTGTSATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHVQEGQGT 60

QY 61 EVVQVLLGLVVEDRLITNCFFPPQHTEDDADFDEVQYQEMMRSLRHVNIDHLHVGWY 120

DB 61 EVVQVLLGLVVEDRLITNCFFPPQHTEDDADFDEVQYQEMMRSLRHVNIDHLHVGWY 120

QY 121 QSTYGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEK 180

DB 121 QSTYGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEK 180

QY 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSAAVADKHELLSLASSNHLG 240

DB 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSAAVADKHELLSLASSNHLG 240

QY 241 KNQLQLMDRVDEMSSQDIVKNTYMRNTSKQQQKHQYQRRQENMQRSRGPPLPEED 300

DB 241 KNQLQLMDRVDEMSSQDIVKNTYMRNTSKQQQKHQYQRRQENMQRSRGPPLPEED 300

QY 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

DB 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

RESULT 3

ADJ68541

ID ADJ68541 standard; protein; 352 AA.

XX AC ADJ68541;

XX XX 06-MAY-2004 (first entry)

XX DT Human heat mitochondrial protein as a therapeutic target SeqID347.

XX DE Mitochondrial; human; screening assay; diabetes mellitus;

XX KW Huntington's disease; osteoarthritis;

XX KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;  
KW osteopathic; ophthalmological; cytostatic.  
XX  
OS Homo sapiens.  
XX  
FN WO2003087768-A2.  
XX  
PD 23-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010870.  
XX  
XX 12-APR-2002; 2002US-0372843P.  
PR 17-JUN-2002; 2002US-0389987P.  
PR 20-SEP-2002; 2002US-0412418P.  
XX  
XX (MITO-) MITOKOR.  
PA (BUCK-) BUCK INST AGE RES.  
XX  
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
PI Warnock DB;  
XX WPI; 2003-845369/78.  
DR  
XX  
XX Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.  
XX  
XX Claim 1; SEQ ID NO 347; 180pp; English.  
XX  
XX This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, nontropic, antidiabetic,  
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
CC cytostatic activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.  
XX  
SQ Sequence 352 AA;

Query Match 99.8%; Score 1808; DB 7; Length 352;  
Best Local Similarity 99.7%; Pred. No. 7.9e-159;  
Matches 351; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MASRKEGTGSTATSSSTAGAGKGGKGGSGDSAVKQVQIDGLVWLKIKHYEEGGGT 60  
Db 1 MASRKEGTGSTATSSSTAGAGKGGKGGSGDSAVKQVQIDGLVWLKIKHYEEGGGT 60  
Qy 61 EVVQGVLLGLVVDRLTNCPPFPQHTDDADFEVQYQNMRRSLRHVNIDHLHVGWY 120  
Db 61 EVVQGVLLGLVVDRLTNCPPFPQHTDDADFEVQYQNMRRSLRHVNIDHLHVGWY 120  
Qy 121 QSTYIGSPVTRALLDSQSYQHATEESVVLVYDPIKTAQGSLSLKAVRLTLPKMEVCKEK 180  
Db 121 QSTYIGSPVTRALLDSQSYQHATEESVVLVYDPIKTAQGSLSLKAVRLTLPKMEVCKEK 180  
Qy 181 DFSPEALKKANITFEYMFEEPIVINKSHLINVLWLEKEKSAVADKHELLSLSSNHLG 240  
Db 181 DFSPEALKKANITFEYMFEEPIVINKSHLINVLWLEKEKSAVADKHELLSLSSNHLG 240  
Qy 241 KNLQLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQQRQENMQRQSRGEPPLPEED 300  
Db 241 KNLQLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQQRQENMQRQSRGEPPLPEED 300

Qy 301 LSKLKFPPQPPARMDSLLIAGQINTYCNQIKETFAQNGLKLFMAQALQEYNN 352  
Db 301 LSKLKFPPQPPARMDSLLIAGQINTYCNQIKETFAQNGLKLFMAQALQEYNN 352  
RESULT 4  
ADC31312  
ID ADC31312 standard; protein; 368 AA.  
XX  
XX ADC31312;  
AC  
DT 18-DEC-2003 (first entry)  
DE  
DE Human novel polypeptide sequence, SEQ ID NO:1394.  
XX  
XX Human; diagnostic; drug screening; forensics; gene mapping;  
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KW ulcers; osteoporosis; autoimmune disease; cancer;  
KW molecular weight marker; food supplement; antiparkinsonian; nontropic;  
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnery;  
KW anticulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KW gene therapy; chromosome 8p22.  
XX  
XX Homo sapiens.  
OS  
XX WO2003029271-A2.  
PN  
XX 10-APR-2003.  
PD  
XX 24-SEP-2002; 2002WO-US030474.  
PF  
XX 24-SEP-2001; 2001US-0324631P.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX WPI; 2003-371981/35.  
DR N-PSDB; ADC30341.  
XX  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
XX Claim 20; SEQ ID NO 1394; 1185pp; English.  
XX  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's





QY 301 LSKLFKPPPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQYNN 352  
 DB 301 LSKLFKPPPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQYNN 352

RESULT 6  
 ADE29059  
 ID ADE29059 standard; protein; 347 AA.  
 XX AC ADE29059;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human EIF3S3 selected interacting domain protein - SEQ ID 45.  
 XX KW anti-HIV; SID; selected interacting domain; HIV; gene therapy; siRNA;  
 KW silencing RNA; HIV-1 integrase; LEDGF; MCM7; HBO1; Snurportin; VBPI;  
 KW Transportin-SR; EIF3S3; human.  
 XX OS Homo sapiens.  
 XX PN WO2003046176-A2.  
 XX PD 05-JUN-2003.  
 XX PF 26-NOV-2002; 2002WO-EP013868.  
 XX PR 26-NOV-2001; 2001US-0333346P.  
 XX PR 31-MAY-2002; 2002US-0385132P.  
 XX PA (HYBR-) HYBRIGENICS.  
 XX PI Legrain P, Rain J, Benarous R, Emiliani S, Berlioz-Torrent C;  
 PI Blot G;  
 XX WPI; 2003-505199/47.  
 XX DR N-PSDB; ADE29036.  
 XX PT New complex between two interacting proteins, useful for screening  
 PT molecules that inhibit human immunodeficiency virus or for preparing a  
 PT medicament for treating HIV-1.  
 XX PS Claim 9; SEQ ID NO 45; 102pp; English.  
 XX CC The invention relates to a novel complex between two interacting proteins  
 CC listed within the specification. The complex of the invention  
 CC demonstrates anti-HIV activity whilst the SID (selected interacting  
 CC domains) and polypeptides may be useful for screening molecules that  
 CC inhibit human immunodeficiency virus (HIV), as well as during gene  
 CC therapy procedures. The siRNAs (silencing RNAs) targeted against cellular  
 CC proteins interacting with HIV-1 integrase, LEDGF, MCM7, HBO1, Snurportin,  
 CC VBPI, Transportin-SR and EIF3S3, are useful in preparing a medicament for  
 CC treating HIV-1. The current sequence is that of the human SID protein of  
 CC the invention.  
 XX SQ Sequence 347 AA;

Query Match 98.7%; Score 1789; DB 7; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-157;  
 Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 EGTGSTATSSSTAGAGKGGSGDSAVKQVQIDGLVVLKIKKHQYBEGQGTVEVQG 65  
 DB 1 EGTGSTATSSSTAGAGKGGSGDSAVKQVQIDGLVVLKIKKHQYBEGQGTVEVQG 60  
 QY 66 VLLGLVWEDRLIETNCFPPQHTEDDADFDEVOYQEMWRSRLRVNIDHLVGVYQSTYY 125  
 DB 61 VLLGLVWEDRLIETNCFPPQHTEDDADFDEVOYQEMWRSRLRVNIDHLVGVYQSTYY 120

QY 126 GSFVTRALLDSQFSYQHAIEBSVLLIYDPIKTAQSLSLKAYRLTPKLMVECKSKDPSPE 185  
 DB 121 GSFVTRALLDSQFSYQHAIEBSVLLIYDPIKTAQSLSLKAYRLTPKLMVECKSKDPSPE 180

QY 186 ALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADADKHELLSLASSNHLGKNLQL 245  
 DB 181 ALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADADKHELLSLASSNHLGKNLQL 240

QY 246 LMDRVDEMSQDIVKYNTYMENTSQQQQKHQYQORROENMQRSGEPPLPEEDLSKLF 305  
 DB 241 LMDRVDEMSQDIVKYNTYMENTSQQQQKHQYQORROENMQRSGEPPLPEEDLSKLF 300

QY 306 KPPOPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQYNN 352  
 DB 301 KPPOPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQYNN 347

RESULT 7  
 ABB08367  
 ID ABB08367 standard; protein; 352 AA.  
 XX AC ABB08367;  
 XX DT 07-MAY-2002 (first entry)  
 XX DE Human cancer antigen eIF3 variant 2 amino acid sequence.  
 XX KW Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 242 /note= "wild-type Asn is replaced by Phe"  
 FT Domain 242 /note= "HLA-2 binding residue"  
 FT Domain 243 /note= "HLA-2 binding residue"  
 FT Domain 244..249 /note= "HLA-2 binding residue"  
 FT Misc-difference 246 /note= "T-cell receptor (TCR) binding domain"  
 FT Misc-difference 247 /note= "wild-type Leu is replaced by Glu"  
 FT Misc-difference 247 /note= "wild-type Met is replaced by Phe"  
 FT Misc-difference 249 /note= "wild-type Arg is replaced by Ala"  
 FT Domain 250 /note= "HLA-2 binding residue"  
 FT WO200192307-A2.  
 XX 06-DEC-2001.  
 XX 30-MAY-2001; 2001WO-US017456.  
 XX 31-MAY-2000; 2000US-0209391P.  
 XX 17-AUG-2000; 2000US-0226258P.  
 XX 20-DEC-2000; 2000US-0257008P.  
 XX (GENZ ) GENZYME CORP.  
 XX Nicolette CA;  
 XX WPI; 2002-139606/18.  
 XX New therapeutic compounds useful against human ovarian cancer, for  
 XX modulating immune response in a subject, and for generating antibodies  
 XX that specifically recognize and bind to these molecules.  
 XX Claim 7; Page; 68pp; English.  
 XX The invention relates to novel therapeutic compounds, that are designed  
 XX to enhance binding to MHC molecules and to enhance immunoregulatory  
 XX properties relative to their natural counterparts. The activity of the

CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.

CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen eIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human cancer antigen eIF3 variant 2 amino acid sequence.

CC Note: This sequence is not present in the specification, but may be  
 CC created from the sequence of the wild-type human cancer antigen eIF3  
 CC sequence given in ABB08360

XX Sequence 352 AA;

Query Match 98.5%; Score 1785; DB 5; Length 352;  
 Best Local Similarity 98.9%; Pred. No. 1.1e-156;  
 Matches 348; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVWLKIKHYQEGQGT 60

Db 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVWLKIKHYQEGQGT 60

Qy 61 EVVQGVLLGLVVEDRLITNCFFPQHTEDDADFEVQYQWEMMRSRLRVNIDHLHVGWY 120

Db 61 EVVQGVLLGLVVEDRLITNCFFPQHTEDDADFEVQYQWEMMRSRLRVNIDHLHVGWY 120

Qy 121 QSTYGGSFVTRALLDSQFSYQHATEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180

Db 121 QSTYGGSFVTRALLDSQFSYQHATEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180

Qy 181 DFSPEALKKANITPEYMFEEVPIVKNSHLINVLWLEKKSADKHELLSLASSNHLG 240

Db 181 DFSPEALKKANITPEYMFEEVPIVKNSHLINVLWLEKKSADKHELLSLASSNHLG 240

Qy 241 KNLQLLMDRVDEMSQDITVKNTYMRNTSKQQKHQYQRRQENMQRSGRPPLPEED 300

Db 241 KFLQLEFADVDEMSQDITVKNTYMRNTSKQQKHQYQRRQENMQRSGRPPLPEED 300

Qy 301 LSKLFKPPQPPARMDSLLIAGQINTYCONIKFTAQNLGKLFMAQALQEVNN 352

Db 301 LSKLFKPPQPPARMDSLLIAGQINTYCONIKFTAQNLGKLFMAQALQEVNN 352

RESULT 8

ABB08368

ID ABB08368 standard; protein; 352 AA.

XX AC ABB08368;

DT 07-MAY-2002 (first entry)

XX Human cancer antigen eIF3 variant 3 amino acid sequence.

XX Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 242 /note= "wild-type Asn is replaced by Phe"

FT Domain 242 /note= "HLA-2 binding residue"

FT Domain 243 /note= "HLA-2 binding residue"

FT Domain 244. .249

FT Domain 244. .249

FT Misc-difference 244 /note= "T-cell receptor (TCR) binding domain"

FT Misc-difference 245 /note= "wild-type Gln is replaced by Trp"

FT Misc-difference 246 /note= "wild-type Leu is replaced by Phe"

FT Misc-difference 247 /note= "wild-type Leu is replaced by Glu"

FT Misc-difference 249 /note= "wild-type Met is replaced by Ile"

FT Domain 250 /note= "wild-type Arg is replaced by Ile"

XX WO200192307-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017456.

PR 31-MAY-2000; 2000US-0209391P.

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

XX (GENZ ) GENZYME CORP.

XX Nicolette CA;

XX WPI; 2002-139606/18.

XX New therapeutic compounds useful against human ovarian cancer, for  
 PT modulating immune response in a subject, and for generating antibodies  
 PT that specifically recognize and bind to these molecules.  
 XX Claim 8; Page; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen eIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human cancer antigen eIF3 variant 3 amino acid sequence.  
 CC Note: This sequence is not present in the specification, but may be  
 CC created from the sequence of the wild-type human cancer antigen eIF3  
 CC sequence given in ABB08360

XX Sequence 352 AA;

Query Match 97.8%; Score 1773; DB 5; Length 352;

Best Local Similarity 98.3%; Pred. No. 1.4e-155;

Matches 346; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVWLKIKHYQEGQGT 60

Db 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVWLKIKHYQEGQGT 60

Qy 61 EVVQGVLLGLVVEDRLITNCFFPQHTEDDADFEVQYQWEMMRSRLRVNIDHLHVGWY 120

Db 61 EVVQGVLLGLVVEDRLITNCFFPQHTEDDADFEVQYQWEMMRSRLRVNIDHLHVGWY 120

Qy 121 QSTYGGSFVTRALLDSQFSYQHATEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180

Db 121 QSTYGGSFVTRALLDSQFSYQHATEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180

Db 121 QSTYGSFVTRALLDSQFSYQHAIESVVLVYDPIKTAQGSLSLKAYRLTPKLMVEVCCKE 180  
Qy 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSAVADKHELLSLASSNHLG 240  
Db 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSAVADKHELLSLASSNHLG 240  
Qy 241 KNLQLLMDRVDEMSSQDIVKYNTYMRNTSKQQQKHQYQQRQENMQQRSGEPPLPEED 300  
Db 241 KFLWFEIDVDEMSSQDIVKYNTYMRNTSKQQQKHQYQQRQENMQQRSGEPPLPEED 300  
Qy 301 LSKLFPKPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352  
Db 301 LSKLFPKPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

RESULT 9

ABB08369  
ID ABB08369 standard; protein; 352 AA.

AC ABB08369;

DT 07-MAY-2002 (first entry)

XX Human cancer antigen eIF3 variant 4 amino acid sequence.

DE Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
KW ovarian cancer; MHC; cytosstatic; immunomodulator; immune effector cell;  
KW anti-cancer; vaccine.  
XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 242 /note= "wild-type Asn is replaced by Phe"

FT Domain 242 /note= "HLA-2 binding residue"

FT Domain 243 /note= "HLA-2 binding residue"

FT Domain 244..249 /note= "T-cell receptor (TCR) binding domain"

FT Misc-difference 244 /note= "wild-type Gln is replaced by Ser"

FT Misc-difference 245 /note= "wild-type Leu is replaced by Tyr"

FT Misc-difference 246 /note= "wild-type Leu is replaced by Asp"

FT Misc-difference 247 /note= "wild-type Met is replaced by Leu"

FT Misc-difference 248 /note= "wild-type Asp is replaced by Phe"

FT Misc-difference 249 /note= "wild-type Arg is replaced by Val"

FT Domain 250 /note= "HLA-2 binding residue"

FT Domain 250 /note= "HLA-2 binding residue"

XX WO200192307-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017456.

XX 31-MAY-2000; 2000US-0209391P.

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

XX (GENZ ) GENZYME CORP.

XX Nicolette CA;

XX WPI; 2002-139606/18.

XX New therapeutic compounds useful against human ovarian cancer, for

FT modulating immune response in a subject, and for generating antibodies

PT that specifically recognize and bind to these molecules.

XX Claim 9; Page; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory  
CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.

CC compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterised by expression of antigen eIF3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detecting genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents the human cancer antigen eIF3 variant 4 amino acid sequence.  
CC Note: This sequence is not present in the specification, but may be  
CC created from the sequence of the wild-type human cancer antigen eIF3  
CC sequence given in ABB08360

XX Sequence 352 AA;

Query Match 97.4%; Score 1765; DB 5; Length 352;

Best Local Similarity 98.0%; Pred. No. 7.7e-155;

Matches 345; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MASRKEGTGSTATSSSSSTAGAGKGGKGGSGDSAVKQVQIDGLVWLKIKHYQEGOGT 60

Db 1 MASRKEGTGSTATSSSSSTAGAGKGGKGGSGDSAVKQVQIDGLVWLKIKHYQEGOGT 60

Qy 61 EVQGVLLGLVVEDRLITNCFPPTHEDDADFEVQYQEMMRSLRHVNIDHLHVGWY 120

Db 61 EVQGVLLGLVVEDRLITNCFPPTHEDDADFEVQYQEMMRSLRHVNIDHLHVGWY 120

Qy 121 QSTYGSFVTRALLDSQFSYQHAIESVVLVYDPIKTAQGSLSLKAYRLTPKLMVEVCCKE 180

Db 121 QSTYGSFVTRALLDSQFSYQHAIESVVLVYDPIKTAQGSLSLKAYRLTPKLMVEVCCKE 180

Qy 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSADKHELLSLASSNHLG 240

Db 181 DFPSEALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSADKHELLSLASSNHLG 240

Qy 241 KNLQLLMDRVDEMSSQDIVKYNTYMRNTSKQQQKHQYQQRQENMQQRSGEPPLPEED 300

Db 241 KFLSYDLFWVDDEMSSQDIVKYNTYMRNTSKQQQKHQYQQRQENMQQRSGEPPLPEED 300

Qy 301 LSKLFPKPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

Db 301 LSKLFPKPQPPARMDSLLIAGQINTYCNITKEFTAQNLGKLFMAQALQEYNN 352

RESULT 10

ADE29060

ID ADE29060 standard; protein; 302 AA.

XX AC ADE29060;

XX 29-JAN-2004 (first entry)

DT Human eIF3S3 selected interacting domain protein - SEQ ID 46.

XX anti-HIV; SID; selected interacting domain; HIV; gene therapy; siRNA;

XX silencing RNA; HIV-1 integrase; LEDGF; MCW7; HB01; Smurportin; VB1;

XX Transportin-SR; eIF3S3; human.

XX Homo sapiens.

XX OS

XX WO2003046176-A2.

FN

XX PD 05-JUN-2003.  
XX PF 26-NOV-2002; 2002WO-EP013868.  
XX XX 26-NOV-2001; 2001US-0333346P.  
XX PR 31-MAY-2002; 2002US-0385132P.  
XX PA (HYBR-) HYBRIGENICS.  
XX PI Legrain P, Rain J, Benarous R, Emiliani S, Berlioz-Torrent C;  
PI Blot G;  
XX WPI; 2003-505199/47.  
XX DR N-PSDB; ADE29037.  
XX DR New complex between two interacting proteins, useful for screening a  
PT molecules that inhibit human immunodeficiency virus or for preparing a  
PT medicament for treating HIV-1.  
XX PS Claim 9; SEQ ID NO 46; 102pp; English.  
XX PS  
XX CC The invention relates to a novel complex between two interacting proteins  
CC listed within the specification. The complex of the invention  
CC demonstrates anti-HIV activity whilst the SID (selected interacting  
CC domains) and polypeptides may be useful for screening molecules that  
CC inhibit human immunodeficiency virus (HIV), as well as during gene  
CC therapy procedures. The siRNAs (silencing RNAs) targeted against cellular  
CC proteins interacting with HIV-1 integrase, LEDGF, MCM7, HB01, Snurportin,  
CC VBP1, Transportin-SR and E1F3s3, are useful in preparing a medicament for  
CC treating HIV-1. The current sequence is that of the human SID protein of  
CC the invention.  
XX SQ Sequence 302 AA;  
Query Match 85.7%; Score 1552; DB 7; Length 302;  
Best Local Similarity 100.0%; Pred. No. 3.6e-135;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KEGTGSTATSSSTAGAGKGGSGDSAVKQVQIDGLVLIKHVYQEGQGVVQ 64  
DB 1 KEGTGSTATSSSTAGAGKGGSGDSAVKQVQIDGLVLIKHVYQEGQGVVQ 60  
QY 65 GVLLGLVVEDRLTNCFFPQHTDDADDFVQYQMEMMRSRHNIDHLHVGWYQSTY 124  
DB 61 GVLLGLVVEDRLTNCFFPQHTDDADDFVQYQMEMMRSRHNIDHLHVGWYQSTY 120  
QY 125 YGSPVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEDFSP 184  
DB 121 YGSPVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEDFSP 180  
QY 185 EALKKANITFEYMFEEVPIVKNSHLINVLMWELEKKSADKHELLSLASSNHLGNLQ 244  
DB 181 EALKKANITFEYMFEEVPIVKNSHLINVLMWELEKKSADKHELLSLASSNHLGNLQ 240  
QY 245 LLMDRVDEMSQDIKVNNTYMRNTSKQQQKHQYQORRQENMQRQGRGEPPLPEEDLSKL 304  
DB 241 LLMDRVDEMSQDIKVNNTYMRNTSKQQQKHQYQORRQENMQRQGRGEPPLPEEDLSKL 300  
QY 305 FK 306  
DB 301 FK 302  
RESULT 11  
AAB56775  
ID AAB56775 standard; protein; 256 AA.  
XX AAB56775;  
AC AAB56775;  
XX 13-MAR-2001 (first entry)  
DT Human prostate cancer antigen protein sequence SEQ ID NO:1353.  
DE

XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
KW wound; infectious disease.  
XX OS Homo sapiens.  
XX XX WO2000055174-A1.  
XX PN 21-SEP-2000.  
XX PD 08-MAR-2000; 2000WO-US005988.  
XX PF 12-MAR-1999; 99US-0124270P.  
XX PR (HUMA-) HUMAN GENOME SCI INC.  
XX PA (ROSE/) ROSEN C A.  
XX PI Rosen CA, Ruben SM;  
XX WPI; 2000-587513/55.  
XX DR N-PSDB; AAF15978.  
XX PT Prostate cancer associated gene sequences, referred to as prostate cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as prostate cancer.  
XX PS Claim 11; Page 1777-1778; 2338pp; English.  
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated  
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
CC The prostate cancer antigens can have neuroprotective, cytostatic,  
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAB57303 represent sequences used in the exemplification of the present  
CC invention  
XX SQ Sequence 256 AA;  
Query Match 70.4%; Score 1275; DB 3; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.4e-109;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 107 LRHVNIDHLHVGWYQSTYVGFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKA 166  
DB 11 LRHVNIDHLHVGWYQSTYVGFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKA 70  
QY 167 YRLTPKLMVEVCKEDFSPKALKNITFEYMFEEVPIVKNSHLINVLMWELEKKSAD 226  
DB 71 YRLTPKLMVEVCKEDFSPKALKNITFEYMFEEVPIVKNSHLINVLMWELEKKSAD 130  
QY 227 KHELLSLASSNHLGNLQLLMDRVDEMSQDIKVNNTYMRNTSKQQQKHQYQORRQENMQ 286  
DB 131 KHELLSLASSNHLGNLQLLMDRVDEMSQDIKVNNTYMRNTSKQQQKHQYQORRQENMQ 190  
QY 287 QRSRGEPPLPEEDLSKLFKPPQPPARMDSLLIAGQINTYQNTKEFTAQNLGLFMAQA 346  
DB 191 QRSRGEPPLPEEDLSKLFKPPQPPARMDSLLIAGQINTYQNTKEFTAQNLGLFMAQA 250  
QY 347 LQEYNN 352  
DB 251 LQEYNN 256

RESULT 12  
 ABP42329  
 ID ABP42329 standard; protein; 202 AA.  
 XX  
 AC ABP42329;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HNOB83, SEQ ID NO:3461.  
 XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001WO-US018569.  
 XX  
 PR 07-JUN-2000; 2000US-0209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI; 2002-147878/19.  
 DR N-PSDB; ABQ55406.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 PS Claim 11; SEQ ID NO 3461; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 202 AA;  
 Query Match 56.6%; Score 1026; DB 5; Length 202;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-86;  
 Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 4 RKEGTGSTATSSSTAGAGKGGKGGGSAVKQVQIDGLVVLKIKHKHYOEGQTEWV 63  
 Db 1 RKEGTGSTATSSSTAGAGKGGKGGGSAVKQVQIDGLVVLKIKHKHYOEGQTEWV 60  
 Qy 64 QGVLLGLVVEDRLITNCFFPQHTEDDADFDEYQYQWEMMRSLSRHVNIDHLHVGWYQST 123  
 Db 61 QGVLLGLVVEDRLITNCFFPQHTEDDADFDEYQYQWEMMRSLSRHVNIDHLHVGWYQST 120  
 Qy 124 YGGSFVTRALLDSQFSQHAIESVLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEKDFS 183  
 Db 121 YGGSFVTRALLDSQFSQHAIESVLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKEKDFS 180  
 Qy 184 PEALKKANITFEYMFEEVP 202  
 Db 181 PEALKKANITFEYMFEEVP 199  
 RESULT 13  
 ABB64255  
 ID ABB64255 standard; protein; 307 AA.  
 XX  
 AC ABB64255;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 19557.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL08358.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 19557; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL13051), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-  
 CC ABF72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

3.



PR 23-JUL-1999;	99US-0145224P.
PR 26-JUL-1999;	99US-0145276P.
PR 27-JUL-1999;	99US-0145913P.
PR 27-JUL-1999;	99US-0145918P.
PR 27-JUL-1999;	99US-0145918P.
PR 28-JUL-1999;	99US-0145951P.
PR 02-AUG-1999;	99US-0146386P.
PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 04-AUG-1999;	99US-0147302P.
PR 05-AUG-1999;	99US-0147192P.
PR 05-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 06-AUG-1999;	99US-0147416P.
PR 09-AUG-1999;	99US-0147493P.
PR 09-AUG-1999;	99US-0147935P.
PR 10-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 16-AUG-1999;	99US-0149368P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149929P.
PR 23-AUG-1999;	99US-0149902P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151066P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 20-SEP-1999;	99US-0154779P.
PR 22-SEP-1999;	99US-0155139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155659P.
PR 28-SEP-1999;	99US-0156458P.
PR 29-SEP-1999;	99US-0156596P.
PR 04-OCT-1999;	99US-0157117P.
PR 05-OCT-1999;	99US-0157753P.
PR 06-OCT-1999;	99US-0157865P.
PR 07-OCT-1999;	99US-0158029P.
PR 08-OCT-1999;	99US-0158232P.
PR 12-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159294P.
PR 13-OCT-1999;	99US-0159295P.
PR 14-OCT-1999;	99US-0159329P.
PR 14-OCT-1999;	99US-0159330P.
PR 14-OCT-1999;	99US-0159331P.
PR 14-OCT-1999;	99US-0159637P.
PR 18-OCT-1999;	99US-0159638P.
PR 21-OCT-1999;	99US-0159584P.
PR 21-OCT-1999;	99US-0160741P.
PR 21-OCT-1999;	99US-0160767P.
PR 21-OCT-1999;	99US-0160768P.
PR 21-OCT-1999;	99US-0160770P.
PR 21-OCT-1999;	99US-0160814P.
PR 21-OCT-1999;	99US-0160815P.
PR 22-OCT-1999;	99US-0160980P.
PR 22-OCT-1999;	99US-0160981P.

PR 22-OCT-1999;	99US-0160989P.
PR 25-OCT-1999;	99US-0161404P.
PR 25-OCT-1999;	99US-0161405P.
PR 25-OCT-1999;	99US-0161406P.
PR 26-OCT-1999;	99US-0161359P.
PR 26-OCT-1999;	99US-0161360P.
PR 26-OCT-1999;	99US-0161361P.
PR 28-OCT-1999;	99US-0161920P.
PR 28-OCT-1999;	99US-0161992P.
PR 28-OCT-1999;	99US-0161993P.
PR 29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity 34.6%; Score 627.5; DB 3; Length 334;

Mismatches 136; Conservative 67; Mismatches 114; Indels 9; Gaps 5;

Qy 31	SGDSAV---KQVQIDGLAVLKIKHYQEBGQGTGVVQGVLLGLVVEDRLLEITNCPPPPQ 86
Db 10	SKDEAVAPLRVVQIEGLAVLKIKHCKE--FSPTLVTGQLGLDVGSLVETVNCFFPPV 67
Qy 87	HTED-DADFDVQVQYQEMMESLRHVNIIDLHLVGVQSTYYTGYGSFVTRALLDSQFSYQHAIE 145
Db 68	RDDDEIEADGANYQLEMMRCLREVNVNNTVGVYQSTVGLSYQTVLIEFTFMYQENIK 127
Qy 146	ESVVLIIYDIPTAQSLSLKAYRLTPKLMVCKEKQFSPALKKCANITFYMFEEVPIVI 205
Db 128	RCVCIYDPSKADLVGLKALKLSDFSMBELYRGNGFTGEKLRKKNFSWMDIPEEIDIKV 187
Qy 206	KNSHLINVLWLELEKKSADAD-KHELILLSASSNHLGNLQLLMDRVDEMSSQDIYKNTYM 264
Db 188	SNSALVSFAFTELETDTPVSGDYDRLHSTTTPPLENNMEFLIKCMDLDSMEQOKFQYY 247
Qy 265	RNTSKQOQKHQYQORROQENMORQSRGEPPLPEEDLSK-LFKPQPQPARMDSLIIAGQI 323
Db 248	RNLSRQQAQQAOWLQKRTENMARKSAGEEPLPEEDPSNPFKAIPESRLESFLITNOV 307
Qy 324	NTYQNIKFTPAQNLGKLFMAQALQE 349
Db 308	SNFCGQINGVAGQNFSLYLTALHD 333

RESULT 15

AAG51721

ID AAG51721 standard; protein; 334 AA.

AC AAG51721,

XX 18-OCT-2000 (first entry)

DT Arabidopsis thaliana protein fragment SEQ ID NO: 65671.

DE Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 23-MAR-1999; 99US-0123548P.

PR 25-MAR-1999; 99US-0125788P.

PR 29-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0126785P.

PR 06-APR-1999; 99US-0127462P.

PR 08-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132488P.  
 PR 06-MAY-1999; 99US-0132488P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0132863P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135623P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0144005P.  
 PR 15-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144684P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 26-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 02-AUG-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 03-AUG-1999; 99US-0146389P.  
 PR 04-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 05-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 06-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 11-AUG-1999; 99US-0148171P.  
 PR 12-AUG-1999; 99US-0148315P.  
 PR 13-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 16-AUG-1999; 99US-0148684P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149175P.  
 PR 20-AUG-1999; 99US-0149428P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999;		99US-0159330P.
PR 14-OCT-1999;		99US-0159331P.
PR 14-OCT-1999;		99US-0159637P.
PR 14-OCT-1999;		99US-0159638P.
PR 18-OCT-1999;		99US-0159584P.
PR 21-OCT-1999;		99US-0160741P.
PR 21-OCT-1999;		99US-0160767P.
PR 21-OCT-1999;		99US-0160768P.
PR 21-OCT-1999;		99US-0160770P.
PR 21-OCT-1999;		99US-0160814P.
PR 21-OCT-1999;		99US-0160815P.
PR 22-OCT-1999;		99US-0160980P.
PR 22-OCT-1999;		99US-0160981P.
PR 22-OCT-1999;		99US-0160989P.
PR 25-OCT-1999;		99US-0161404P.
PR 25-OCT-1999;		99US-0161405P.
PR 25-OCT-1999;		99US-0161406P.
PR 26-OCT-1999;		99US-0161359P.
PR 26-OCT-1999;		99US-0161360P.
PR 26-OCT-1999;		99US-0161361P.
PR 28-OCT-1999;		99US-0161920P.
PR 28-OCT-1999;		99US-0161992P.
PR 28-OCT-1999;		99US-0161993P.
PR 29-OCT-1999;		99US-0162142P.
Query Match		34.6%; Score 627.5; DB 3; Length 334;
Best Local Similarity		41.7%; Pred. No. 2.5e-49;
Matches 136; Conservative 67; Mismatches 114; Indels 9; Gaps 5;		
Qy	31	SGDSAV-----KQVQIDGLVVLKIIKHQBEGQGTGVQGVLLGLVVEDRLIEITNCFFPPQ 86
Db	10	SKDEAVAPPLRVVQIEGLAVLKIHKCKE--FSPTLVGTQLGLDVGSLVLEVTNCFFPPV 67
Qy	87	HTED-DADFDVQYQMMRSLRHVNIDHLHVGWYQSTYGSFVTRALLDSQFSYQHAIE 145
Db	68	RDDDEEIEADGANTQLEMRCLEVNVDNNTVGYQSTVLGSYQTVLIEFTFMNYQENIK 127
Qy	146	ESVVLIVDPIKTAQGSLSLKAYRLTPKLMVEYCKEKFSPALKKANITFEYMFEEVPVI 205
Db	128	RCVCIIVDPSKADLGVLAKALKLSDSFMELYRGNGFTGKLRKFNFSWMDIPEEIPKV 187
Qy	206	KNSHLINVLWMELEKKSAAVAD-KHELLSLASSNHLGKNLQLLMDRVDEMSSQDIVKYNTYM 264
Db	188	SNSALVSFAFTELETPTVPSQGDYDLRLHSSTTFLENNMEFLIKCMDLDSMEQKQFQYY 247
Qy	265	RNTSKQQQKHQYQRRQENNRQSRGEPPLPEEDLSK-LFKPPQPPARMDSLLIAGQI 323
Db	248	RNLSRQQAQQAAWLQKARTENMARKSAGEEPLPEEDPSNPFIKAIPEPSRLESFLITNQV 307
Qy	324	NTYCQNIKEFTAQNLGKLFWAQAQAE 349
Db	308	SNFCQINGVAGQNFSLRLYLTALHD 333

Search completed: December 30, 2004, 20:42:53  
Job time : 414.519 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:35:11 ; Search time 344.907 Seconds  
(without alignments)  
367.126 Million cell updates/sec

Title: US-10-017-327-2

Perfect score: 1812

Sequence: 1 MASRKEGTGSTATSSSTAG.....FTAQNLGKLFMAQALQEYNN 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1812	100.0	352	11	US-09-870-216C-2
2	1812	100.0	352	13	US-10-017-327-2
3	1808	99.8	352	16	US-10-408-765A-347
4	1275	70.4	256	9	US-09-925-300-1353
5	1026	56.6	202	15	US-10-264-049-3461
6	666.5	36.8	347	16	US-10-437-963-115552
7	660.5	36.5	398	15	US-10-425-114-59943
8	649.5	35.8	382	15	US-10-425-114-59123
9	628.5	34.7	369	17	US-10-425-115-298023
10	526	29.0	288	16	US-10-767-701-45306
11	482	26.6	364	17	US-10-425-115-317552
12	482	26.6	386	15	US-10-425-114-67030
13	466	25.7	93	15	US-10-424-599-256678

14	427	23.6	252	17	US-10-425-115-310908
15	392	21.6	209	15	US-10-424-599-152171
16	374.5	20.7	206	15	US-10-424-599-152172
17	304	16.8	59	15	US-10-264-049-2708
18	183	10.1	310	17	US-10-842-740-30
19	183	10.1	315	9	US-09-925-300-1238
20	180.5	10.0	306	14	US-10-369-493-21879
21	176.5	9.7	312	14	US-10-032-585-7305
22	168.5	9.3	307	16	US-10-437-963-183249
23	168.5	9.3	307	16	US-10-437-963-183252
24	187.5	9.2	332	15	US-10-320-797-3130
25	166.5	9.2	115	16	US-10-437-963-137948
26	166.5	9.2	307	17	US-10-425-115-314579
27	166.5	9.2	307	17	US-10-425-115-314582
28	166.5	9.2	307	17	US-10-425-115-314586
29	166.5	9.2	308	14	US-10-369-493-2249
30	182.5	9.0	309	15	US-10-425-114-66549
31	162	8.9	312	14	US-10-369-493-5361
32	161.5	8.9	375	15	US-10-424-599-147253
33	159	8.8	310	15	US-10-424-599-169761
34	158.5	8.7	307	16	US-10-767-701-42270
35	158.5	8.7	357	10	US-09-809-391-755
36	158.5	8.7	357	10	US-09-882-171-755
37	158.5	8.7	357	14	US-10-164-861-755
38	157	8.7	356	17	US-10-739-930-7136
39	156.5	8.6	308	14	US-10-302-267-110
40	156.5	8.6	308	15	US-10-412-699B-768
41	154.5	8.5	208	17	US-10-425-115-270116
42	154.5	8.5	305	14	US-10-128-714-3492
43	154	8.5	375	15	US-10-424-599-185555
44	152	8.4	319	14	US-10-369-493-5756
45	151.5	8.4	335	14	US-10-128-714-8492

ALIGNMENTS

RESULT 1

US-09-870-216C-2  
; Sequence 2, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-216C-2

Query Match	100.0%;	Score 1812;	DB 11;	Length 352;
Best Local Similarity	100.0%;	Pred. No. 9.8e-155;		
Matches	352;	Conservative	0;	Mismatches 0;
Indels	0;	Gaps	0;	
Qy	1	MASRKEGTGSTATSSSTAGAGKGGSGS	AVKQVQIDGLVLVLKIKHYEGGGT	60
Db	1	MASRKEGTGSTATSSSTAGAGKGGSGS	AVKQVQIDGLVLVLKIKHYEGGGT	60
Qy	61	EVQGVLLGLVDERLITNCFFPQHTDAD	DEVOYQMEMRSLRHVNIHLHVGWY	120
Db	61	EVQGVLLGLVDERLITNCFFPQHTDAD	DEVOYQMEMRSLRHVNIHLHVGWY	120

QY 121 QSTYGSFVTRALLDSQFSYQHAIESVVLIIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180  
DB 121 QSTYGSFVTRALLDSQFSYQHAIESVVLIIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180  
QY 181 DFSPEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADKHELLSLASSNHLG 240  
DB 181 DFSPEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADKHELLSLASSNHLG 240  
QY 241 KNQLLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQORROENMQRSGEPPLPEED 300  
DB 241 KNQLLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQORROENMQRSGEPPLPEED 300  
QY 301 LSKLFPQPAPPARMDSLLIAGQINTYCNIKFTAQNLGKLFMAQALQEYNN 352  
DB 301 LSKLFPQPAPPARMDSLLIAGQINTYCNIKFTAQNLGKLFMAQALQEYNN 352

## RESULT 2

US-10-017-327-2  
; Sequence 2, Application US/10017327  
; Publication No. US20020155471A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND  
; FILE OF INVENTION: METHODS FOR USING SAME  
; FILE REFERENCE: GZ 2101.20  
; CURRENT APPLICATION NUMBER: US/10/017,327  
; CURRENT FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-017-327-2

Query Match 100.0%; Score 1812; DB 13; Length 352;  
Best Local Similarity 100.0%; Pred. No. 9.8e-155;  
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHQYQEGQGT 60  
DB 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHQYQEGQGT 60  
QY 61 EVVQGVLLGLVVEDRLITNCFFPQHTDDADDFDEVQYQMEMMRSRLRVNIDHLHVGY 120  
DB 61 EVVQGVLLGLVVEDRLITNCFFPQHTDDADDFDEVQYQMEMMRSRLRVNIDHLHVGY 120  
QY 121 QSTYGSFVTRALLDSQFSYQHAIESVVLIIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180  
DB 121 QSTYGSFVTRALLDSQFSYQHAIESVVLIIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180  
QY 181 DFSPEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADKHELLSLASSNHLG 240  
DB 181 DFSPEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADKHELLSLASSNHLG 240  
QY 241 KNQLLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQORROENMQRSGEPPLPEED 300  
DB 241 KNQLLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQORROENMQRSGEPPLPEED 300  
QY 301 LSKLFPQPAPPARMDSLLIAGQINTYCNIKFTAQNLGKLFMAQALQEYNN 352  
DB 301 LSKLFPQPAPPARMDSLLIAGQINTYCNIKFTAQNLGKLFMAQALQEYNN 352

## RESULT 3

US-10-408-765A-347  
; Sequence 347, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Ghosh, Eoin D.  
; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 347  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-347

Query Match 99.8%; Score 1808; DB 16; Length 352;  
Best Local Similarity 99.7%; Pred. No. 2.3e-154;  
Matches 351; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHQYQEGQGT 60  
DB 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHQYQEGQGT 60  
QY 61 EVVQGVLLGLVVEDRLITNCFFPQHTDDADDFDEVQYQMEMMRSRLRVNIDHLHVGY 120  
DB 61 EVVQGVLLGLVVEDRLITNCFFPQHTDDADDFDEVQYQMEMMRSRLRVNIDHLHVGY 120  
QY 121 QSTYGSFVTRALLDSQFSYQHAIESVVLIIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180  
DB 121 QSTYGSFVTRALLDSQFSYQHAIESVVLIIYDPIKTAQGSLSLKAYRLTPKLMVEVCKE 180  
QY 181 DFSPEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADKHELLSLASSNHLG 240  
DB 181 DFSPEALKKANITFEYMFEEVPIVKNSHLINVLWMELEKKSADKHELLSLASSNHLG 240  
QY 241 KNQLLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQORROENMQRSGEPPLPEED 300  
DB 241 KNQLLLMDRVDEMDSQDIVKNTYMRNTSKQQQKHQYQORROENMQRSGEPPLPEED 300  
QY 301 LSKLFPQPAPPARMDSLLIAGQINTYCNIKFTAQNLGKLFMAQALQEYNN 352  
DB 301 LSKLFPQPAPPARMDSLLIAGQINTYCNIKFTAQNLGKLFMAQALQEYNN 352

## RESULT 4

US-09-925-300-1353  
; Sequence 1353, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1353  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (110)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1353

Query Match 70.4%; Score 1275; DB 9; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.8e-106;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LRHNIDHLHVGWQSTYYGSGFVTRALDQSFSYQHAIEESVLIYDPIKTAQGSLSKA 166  
Db 11 LRHNIDHLHVGWQSTYYGSGFVTRALDQSFSYQHAIEESVLIYDPIKTAQGSLSKA 70

Qy 167 YRLTPKMEVCKEKFSPSPEALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSAD 226  
Db 71 YRLTPKMEVCKEKFSPSPEALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSAD 130

Qy 227 KHELSSLASSNHLGKNLQLLMDRVDMSQDIVKNTYMRNTSKQQQKHQYQRRQBNM 286  
Db 131 KHELSSLASSNHLGKNLQLLMDRVDMSQDIVKNTYMRNTSKQQQKHQYQRRQBNM 190

Qy 287 QRSRGEPLPEEDLSKLFKPPPPARMDSLLIAGQINTYCNITKEFTQNLGKLFMAQA 346  
Db 191 QRSRGEPLPEEDLSKLFKPPPPARMDSLLIAGQINTYCNITKEFTQNLGKLFMAQA 250

Qy 347 LQEYNN 352  
Db 251 LQEYNN 256

RESULT 5  
US-10-264-049-3461  
; Sequence 3461, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133PI  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: Patent In Ver. 3.1  
; SEQ ID NO 3461  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-3461

Query Match 56.6%; Score 1026; DB 15; Length 202;  
Best Local Similarity 100.0%; Pred. No. 3.8e-84;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVLVLIKHQYEEGGQTEVV 63  
Db 1 RKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVLVLIKHQYEEGGQTEVV 60

Qy 64 QGVLLGLVVEDRLTNCFFPQHTEDDADFEVQYQMEMRSLRHVNIDHLHVGWQST 123  
Db 61 QGVLLGLVVEDRLTNCFFPQHTEDDADFEVQYQMEMRSLRHVNIDHLHVGWQST 120

Qy 124 YGSGFVTRALDQSFSYQHAIEESVLIYDPIKTAQGSLSLKAYRLTPKMEVCKEKDPS 183  
Db 121 YGSGFVTRALDQSFSYQHAIEESVLIYDPIKTAQGSLSLKAYRLTPKMEVCKEKDPS 180

Qy 184 PEALKKANITFEYMFEEVP 202  
Db 181 PEALKKANITFEYMFEEVP 199

RESULT 6  
US-10-437-963-115552  
; Sequence 115552, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 115552  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_19138C.1.pap  
US-10-437-963-115552

Query Match 36.8%; Score 666.5; DB 16; Length 347;  
Best Local Similarity 40.6%; Pred. No. 2.3e-51;  
Matches 141; Conservative 72; Mismatches 119; Indels 15; Gaps 5;

Qy 16 SSTAGAGKGGKGG-----SGDSAVKQVQIDGLVLVLIKHQYEEGGQTEVVQ 65  
Db 2 ANPAAAGPGGARSFLQAVSTVTEAPSLRVVQMEGLAVLKIHKCEEFAPA--LV 59

Qy 66 VLLGLVVEDRLTNCFFPQHTED-DADFEVQYQMEMRSLRHVNIDHLHVGWQST 124  
Db 60 QLLGLDVGSVLEVTNCFPFPMREDEADADGANYQLEMNRCLEVNVDNNTVGVYQ 119

Qy 125 YGSGFVTRALDQSFSYQHAIEESVLIYDPIKTAQGSLSLKAYRLTPKMEVCKEKD 184  
Db 120 LGSFQTVELIETFMNQENIRRCVICIYDPSRQGVGLALKALKUTDSFMDLYRN 179

Qy 185 EALKKANITFEYMFEEVPIVKNSHLINVLWLEKKSADAD-KHELSSLASSNHLGKN 243  
Db 180 EKLEKLSWVDIFEIPIKVSNSALVSATWLEPESPVSQCDFDLKLTAPFMERN 239

Qy 244 QLLMDRVDMSQDIVKNTYMRNTSKQQQKHQYQRRQBNMQRSGEPPLPEEDLSK 303  
Db 240 EFLGCMDDLSEQNKFQYVYVNRVSRQSQQAQWLQKRRQENMARKAAGEERPLPE 299

Qy 304 -LFXPPQPPARMDSLLIAGQINTYCNITKEFTQNLGKLFMAQALQE 349  
Db 300 PIFKPIEPSRLEGVLVTNQISSYCNHNGVAGQNFNRLYLMKALQE 346

RESULT 7  
US-10-425-114-59943  
; Sequence 59943, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 59943  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:



[illegible]

RESULT 10  
US-10-767-701-45306  
; Sequence 45306, Application US/10767701  
; Publication No. US20040172684A1

Query Match	34.7%	Score 628.5	DB 17	Length 369
Best Local Similarity	39.4%	Pred. No. 6.8e-48		
Matches 135	Conservative 68	Mismatches 109	Indels 31	Gaps 5
QY	36	VKVOQIDGLVVLKIIKHYYQEGQGTGVVQVGLGLGVLEDRLEITNCFFPQHTEDD	-ADP	94
Db	28	LRVVQMEGLAVLKIIKHCEEPAPA	-LVTQQLGLDVGSLVETNCFFPFWREEDDEADA	85
QY	95	DEVQYQEMMRRSLRHVNIIDHLHWGQSTYTGSGFVTRALLDSSQSYQHAIEESVVLIVDP	154	
Db	86	DGANYQLEMRRCLREYVNDNTIGWYQSCLLGSFQIVELIETFMNYQESIRRCVCIYDP	145	
QY	155	IKTAQCSLSLKAYRLTPKLMVEVCKE	-----KDPSPALK	188
Db	146	SRSSQVGLAKALKULTDSFMDLRNNGLTGKYNLFSAVHCTSLYKYFLYNDLFLXLRL	205	
QY	189	KANITPEYMFEEVPIVTKSHLNLVLMWELEKKSAVAD	-KHEILLSASSNHLGNQLQLLM	247
Db	206	EKLSLWDFEELIPLKSSALSVAFWKELEPSPVTCQDLRLKLTASTPFMRNLBFLI	265	
QY	248	DRYDENSQDIVKYNTYMYRNTSKQQQKHQYQQRROQENMQRSGRPPLPEEDLSK	-LFK	306
Db	266	GCWDDLSLSEQNKFQYNNRLSRQSQSQQAWLQKROENMARKAAGEPIFEEDPSNPIFK	325	
QY	307	PQPPAPRMSLLIAGQINTYQCNIKEFTAQNLKLFMAQALQE	349	
Db	326	PIPEPSRLEGYLVNTOICSYCHNINGVAGONFRLYLMKALHE	368	

```

RESULT 11
US-10-425-115-317552
; Sequence 317552, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 317552
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_52685C.1.pgp
US-10-425-115-317552

```

```
Db 206 HDILVELPVTIHNHSLTSLHLQLPSS---APKEELKPPPPNLAALQQDPNIPQPLFPNY 262
Qy 231 --LSLASSNHLGKQLQLMDRVDEMSQDIVKYNTYMTNTSKQOQKHQYQORROENMQR 288
Db 263 DSLDLSIDPFLKCTDLLESIEHNHTEINNYQYQSLAREQAKITAWOQKKAENAR 322
Qy 289 QSRGEPPLPEEDSLKFLKPPQPPARMDSLLTAGQINTVYCNKEFTAGNLGKLFMAQA 346
Db 323 TASKQSLLPEDEWQRLFKLPQEPSRLTLLNSRQVEQYSRQVDFGTAGVTSKMAVKS 380

RESULT 13
US-10-424-599-256678
; Sequence 256678, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 256678
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(93)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_73803C.1.pep
US-10-424-599-256678

Query Match 25.7%; Score 466; DB 15; Length 93;
Best Local Similarity 98.9%; Pred. No. 4.3e-34;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIKHYQERGQGT 60
Db 1 MASRKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIKHYQERGQGT 60

Qy 61 EVVQGVLLGLVVDRLRLTNCPEPPOHTEDDAD 93
Db 61 EVVQGVLLGLVVDRLRLTNCPEPPOHTEDDAD 93

RESULT 14
US-10-425-115-310908
; Sequence 310908, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 310908
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(252)
```

```
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46603C.1.pep
US-10-425-115-310908

Query Match 23.6%; Score 427; DB 17; Length 252;
Best Local Similarity 41.3%; Pred. No. 6e-30;
Matches 90; Conservative 50; Mismatches 74; Indels 4; Gaps 3;

Qy 36 VKQVQIDGLVVLKIKHYQERGQGTVEVQVGLGLVVEDRLTNCFPFP- QHTEDDADF 94
Db 27 LRVQMEGLAVLKIHKCEFPAPA--LVTGQLGLDVGSVLEVTNCFPFPIREDDEADA 84
Qy 95 DEVQYQYQENMRSRLRVNIDHLVGVYQSTYYGSFVTRALLDSQFSQHAIEBSVWLIYDP 154
Db 85 DGANYQLEMRCLEEVNDNNTVGMVYQSCLLGSFQTVELIETFMNYQENIRRCVCIVYDP 144
Qy 155 IKTAQSLSLKAYRLTPKLMVECKEDFSPALKKANITFEYMFEEVPVIVIKNSHLINVL 214
Db 145 SRSSQGVLAALKALKLTDSFMDLYRNNGLTGKLRKKLSWVDIFEIPIXVNSALVSAP 204
Qy 215 MMELEKKSADAD-KHELILLSASSNHLGKQLQLMDRVD 251
Db 205 MKELEPEPSPVQCDFXRLKLTSTAPFERNLERNLFGMIGCMD 242

RESULT 15
US-10-424-599-152171
; Sequence 152171, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152171
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108433C.1.pep
US-10-424-599-152171

Query Match 21.6%; Score 392; DB 15; Length 209;
Best Local Similarity 38.7%; Pred. No. 6.6e-27;
Matches 79; Conservative 54; Mismatches 69; Indels 2; Gaps 2;

Qy 148 VVLIYDPIKTAQGSLSLKAYRLTPKLMVECKEDFSPALKKANITFEYMFEEVPVIVKN 207
Db 5 VCIIVDPSRSDQGVLAALKALKLSDSFMELYSNNFTGKLRKNLSWVDIFEIPIKVSN 64
Qy 208 SHLINVLMELEKKSADAD-KHELILLSASSNHLGKQLQLMDRVDMSQDIVKYNTYMRN 266
Db 65 SALISAFMTELEPDPTFTQCDYDRLQLSTSLMERNTFLIECDMDLSLEQKQKQFYYS 124
Qy 267 TSKQOQKHQYQORROENMQRGEPPLPEEDLSK-LFKPPQPPARMDSLLIAGQINT 325
Db 125 LSRQQAQQQAWLQKRAENMARKAAGEPFLPEEDPSNFIFKPLPSPRLSEFLITNQISN 184
Qy 326 YCONIKETAGNLGKLFMAQALQE 349
Db 185 YCNQINGVSGQSFNRLYLMKALHE 208

Search completed: December 30, 2004, 21:00:29
Job time : 346.907 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:30:54 ; Search time 1.97229 Seconds  
(without alignments)  
439.058 Million cell updates/sec

Title: US-10-017-327-3  
Perfect score: 44  
Sequence: 1 FLQLMEPV 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	79.5	379	T14337	RAD23 protein, iso
2	34	77.3	318	E95218	iron-compound ABC
3	34	77.3	318	C98082	hypothetical prote
4	34	77.3	1254	T24897	hypothetical prote
5	33	75.0	113	H86296	F309.1 protein - A
6	33	75.0	318	G89845	hypothetical prote
7	33	75.0	409	S44346	RAD23 protein homo
8	33	75.0	908	S07649	gene col intron 1
9	33	75.0	2764	T13949	neurofibromin - fr
10	33	75.0	2802	T13945	neurofibromin - fr
11	33	75.0	2802	T13947	neurofibromin - fr
12	32	72.7	382	T14336	RAD23 protein, iso
13	32	72.7	480	A92042	xanthine/uracil pe
14	32	72.7	504	T04076	protoporphyriogen
15	32	72.7	536	F86251	hypothetical prote
16	32	72.7	643	A41120	prostaglandin tran
17	32	72.7	687	B64126	transposase homolo
18	32	72.7	1101	S58108	hypothetical prote
19	32	72.7	1510	T16927	hypothetical prote
20	31	70.5	142	C36179	gonadotropin II be
21	31	70.5	142	I50143	gonadotropin II be
22	31	70.5	255	A69939	conserved hypothet
23	31	70.5	342	T29192	hypothetical prote
24	31	70.5	401	T01001	hypothetical prote
25	31	70.5	456	S69677	hypothetical prote
26	31	70.5	539	T10626	reticuline oxidase
27	31	70.5	639	A55019	muscarinic acetyl
28	31	70.5	730	T43317	pgl-1 protein - Ca
29	31	70.5	771	T29177	hypothetical prote

30	31	70.5	921	2	F81294	probable helicase
31	31	70.5	970	2	D59435	Gem-interacting pr
32	31	70.5	1132	2	A35098	MHC class III hist
33	31	70.5	1760	2	E86201	protein F12K11.4 [
34	30	68.2	96	2	E83533	hypothetical prote
35	30	68.2	105	2	S56917	hypothetical prote
36	30	68.2	126	2	S45797	probable membrane
37	30	68.2	142	1	A25800	gonadotropin beta
38	30	68.2	307	2	T29222	hypothetical prote
39	30	68.2	326	2	G75117	diptetide abc tran
40	30	68.2	329	2	G71014	probable dipeptide
41	30	68.2	358	2	T34128	hypothetical prote
42	30	68.2	367	2	F96827	protein F20B17.8 [
43	30	68.2	395	2	AE3356	lipid-A-disacchari
44	30	68.2	541	2	G86151	F22M8.11 protein -
45	30	68.2	541	2	T08456	hypothetical prote

ALIGNMENTS

RESULT 1

T14337  
RAD23 protein, isoform II - carrot  
C:Species: Daucus carota (carrot)  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 16-Aug-2004  
C:Accession: T14337  
R:Sturm, A.; Leinhardt, S.  
Plant J. 13, 815-821, 1998  
A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.  
A:Reference number: Z17989; MUID:98345997; PMID:9681019  
A:Accession: T14337  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-379 <STU>  
A:Cross-references: UNIPROT:O03991; EMBL:Y12014; NID:gl1914684; PIDN:CAA72742.1; PID:gl1914684  
A:Experimental source: subspecies Queen Anne's Lace, isolate W001C  
C:Genetics:  
A:Gene: RAD23-2  
C:Superfamily: ubiquitin homology

Query Match 79.5%; Score 35; DB 2; Length 379;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
|||: |||  
Db 302 FLQLMEPV 310

RESULT 2

E95218  
iron-compound ABC transporter, permease protein SPI870 [imported] - Streptococcus pneumo  
C:Species: Streptococcus pneumoniae  
C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: E95218  
R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umavam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
naon, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: E95218  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-318 <KUR>  
A:Cross-references: UNIPROT:Q97NY1; GB:AE005672; PIDN:AAK75942.1; PID:gl14973373; GSPDB:G  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SPI870

Query Match 77.3%; Score 34; DB 2; Length 318;

Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEP 8  
|||:|:|  
Db 152 FLOQLMDP 159

RESULT 3  
C99082  
hypothetical protein fatC [imported] - Streptococcus pneumoniae (strain R6)  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
R;Accession: C98082  
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 193, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: C98082  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-318 <KUR>  
A;Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AAL00488.1; PID:g15459360; GSPDB:GN  
A;Gene: fatC

Query Match 77.3%; Score 34; DB 2; Length 318;  
Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEP 8  
|||:|:|  
Db 152 FLOQLMDP 159

RESULT 4  
T24897  
hypothetical protein Tl3H5.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T24897  
R;Lightning, J.  
submitted to the EMBL Data Library, October 1995  
A;Reference number: 219950  
A;Accession: T24897  
A;Status: Preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1254 <WIL>  
A;Cross-references: UNIPROT:Q22467; EMBL:Z66524; PIDN:CAA91418.1; GSPDB:GN000020; CESP.TI  
A;Experimental source: clone Tl3H5  
C;Genetics:  
A;Gene: CESP.Tl3H5.2  
A;Map position: 2  
A;Introns: 63/1; 158/2; 200/3; 286/2; 404/3; 437/1; 539/3; 616/1; 688/3; 732/3; 778/3; 8

Query Match 77.3%; Score 34; DB 2; Length 1254;  
Best Local Similarity 77.8%; Pred. No. 61;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 9  
|||:|:|  
Db 622 FLOQLSVEPV 630

RESULT 5  
H86296  
F309.1 protein - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: H86296

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
anssen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: H86296  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-113 <STO>  
A;Cross-references: UNIPROT:Q9SA20; GB:AE005172; NID:g4966345; PIDN:AAD34676.1; GSPDB:GN  
C;Genetics:  
A;Map position: 1

Query Match 75.0%; Score 33; DB 2; Length 113;  
Best Local Similarity 87.5%; Pred. No. 8.2;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOQLMEP 8  
|||:|:|  
Db 34 FLOQLNEP 41

RESULT 6  
G89845  
hypothetical protein SA0689 [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C;Accession: G89845  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: G89845  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-318 <KUR>  
A;Cross-references: UNIPROT:Q99VP0; GB:BA000018; PID:g13700625; PIDN:BAB41922.1; GSPDB:GN  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SA0689

Query Match 75.0%; Score 33; DB 2; Length 318;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEP 8  
|||:|:|  
Db 151 FLOQLMDP 158

RESULT 7  
S44346  
RAD23 protein homolog - human  
N;Alternate names: p58/HR23B; XP-C repair complementing protein  
C;Species: Homo sapiens (man)  
C;Date: 25-Dec-1994 #sequence\_revision 01-Sep-1995 #text\_change 16-Aug-2004  
C;Accession: S44346  
R;Maetani, C.; Sugasawa, K.; Yanagisawa, J.; Sonoyama, T.; Ui, M.; Enomoto, T.; Takio, T  
EMBO J. 13, 1831-1843, 1994  
A;Title: Purification and cloning of a nucleotide excision repair complex involving the  
A;Reference number: S44345; MUID:94222030; PMID:8168482  
A;Accession: S44346  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-409 <MAS>

A;Cross-references: UNIPROT:P54727; EMBL:D21090; NID:g498147; PIDN:BAA04652.1; PID:di005  
C;Species: Drosophila melanogaster  
C;Superfamily: ubiquitin homology  
F;1-80/Domain: ubiquitin homology <UBH>

Query Match 75.0%; Score 33; DB 2; Length 409;  
Best Local Similarity 66.7%; Pred. No. 32;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
Db 324 YLQMLEPV 332  
:|:|:|

# RESULT 8

S07649  
gene coi intron 1 protein - Neurospora crassa mitochondrion  
C;Species: Neurospora crassa  
C;Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 07-Dec-1999  
C;Accession: S07649  
R;Field, D.J.; Sommerfield, A.; Saville, B.J.; Collins, R.A.  
Nucleic Acids Res. 17, 9087-9099, 1989  
A;Title: A group II intron in the Neurospora mitochondrial coi gene: nucleotide sequence  
A;Reference number: S07649; MUID:90067912; PMID:2531370  
A;Accession: S07649  
A;Molecule type: DNA  
A;Residues: 1-908 <FIE>  
A;Cross-references: EMBL:X14669  
A;Experimental source: strain Adipodoume  
C;Genetics:  
A;Genome: mitochondrion  
A;Genetic code: SGC3  
C;Superfamily: yeast mitochondrion ox13 intron 1 protein  
C;Keywords: mitochondrion

Query Match 75.0%; Score 33; DB 2; Length 908;  
Best Local Similarity 75.0%; Pred. No. 73;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQLMEPV 9  
Db 369 IQLWNEPV 376  
:|:|:|

# RESULT 9

T13949  
neurofibromin - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T13949  
R;The, I.; Hannigan, G.E.; Cowley, G.S.; Reginald, S.; Zhong, Y.; Gusella, J.F.; Harihar  
Science 276, 791-794, 1997  
A;Title: Rescue of a Drosophila Nfi mutant phenotype by protein kinase A.  
A;Reference number: Z17826; MUID:97277221; PMID:9115203  
A;Accession: T13949  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-2764 <THE>  
A;Cross-references: UNIPROT:O01399; EMBL:L26502; NID:g1929432; PIDN:AA8589  
C;Genetics:  
A;Gene: NFI  
A;Cross-references: FlyBase:FBgn0015269

Query Match 75.0%; Score 33; DB 2; Length 2764;  
Best Local Similarity 55.6%; Pred. No. 2.3e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
Db 1340 YLQMLEPL 1348  
:|:|:|

# RESULT 10

T13945

neurofibromin - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T13945  
R;Bernards, A.  
submitted to the EMBL Data Library, April 1997  
A;Reference number: Z17824  
A;Accession: T13945  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-2802 <BER>  
A;Cross-references: UNIPROT:O01397; EMBL:L26500; NID:g1929428; PID:g1929429; PIDN:AA8589  
C;Genetics:  
A;Gene: NFI  
A;Cross-references: FlyBase:FBgn0015269  
A;Introns: 19/3; 67/3; 196/1; 243/2; 441/2; 548/3; 733/1; 912/2; 1443/1; 1645/2; 2430/3;

Query Match 75.0%; Score 33; DB 2; Length 2802;  
Best Local Similarity 55.6%; Pred. No. 2.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
Db 1340 YLQMLEPL 1348  
:|:|:|

# RESULT 11

T13947  
neurofibromin - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T13947  
R;The, I.; Hannigan, G.E.; Cowley, G.S.; Reginald, S.; Zhong, Y.; Gusella, J.F.; Harihar  
Science 276, 791-794, 1997  
A;Title: Rescue of a Drosophila Nfi mutant phenotype by protein kinase A.  
A;Reference number: Z17826; MUID:97277221; PMID:9115203  
A;Accession: T13947  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-2802 <THE>  
A;Cross-references: UNIPROT:O01398; EMBL:L26501; NID:g1929430; PID:g1929431; PIDN:AA8589  
C;Genetics:  
A;Gene: NFI  
A;Cross-references: FlyBase:FBgn0015269

Query Match 75.0%; Score 33; DB 2; Length 2802;  
Best Local Similarity 55.6%; Pred. No. 2.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
Db 1340 YLQMLEPL 1348  
:|:|:|

# RESULT 12

T14336  
RAD23 protein, isoform I - carrot  
C;Species: Daucus carota (carrot)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 16-Aug-2004  
C;Accession: T14336  
R;Sturm, A.; Leinhardt, S.  
Plant J. 13, 815-821, 1998  
A;Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.  
A;Reference number: Z17989; MUID:98345997; PMID:9681019  
A;Accession: T14336  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-382 <STU>  
A;Cross-references: UNIPROT:O03990; EMBL:Y12013; NID:g1914682; PIDN:CAA72741.1; PID:g191  
C;Genetics:  
A;Gene: RAD23-1  
C;Superfamily: ubiquitin homology

Query Match 72.7%; Score 32; DB 2; Length 382;  
 Best Local Similarity 66.7%; Pred. No. 49;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9

Db 311 FLQLINEPM 319

#### RESULT 13

A82042  
 xanthine/uracil permease family protein VC2712 [imported] - Vibrio cholerae (strain N169)  
 C;Species: Vibrio cholerae  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C;Accession: A82042  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301  
 A;Accession: A82042  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-480 <HEI>  
 A;Cross-references: UNIPROT:Q9KMW0; GB:AE004337; GB:AE003852; NID:9657307; PIDN:AAF9589  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC2712  
 A;Map position: 1  
 C;Superfamily: hypothetical protein b2882

Query Match 72.7%; Score 32; DB 2; Length 480;  
 Best Local Similarity 66.7%; Pred. No. 62;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9

Db 371 FVQLIPEPV 379

#### RESULT 14

T04076  
 protoporphyrinogen oxidase (EC 1.3.3.4) IX, mitochondrial [validated] - common tobacco  
 C;Species: Nicotiana tabacum (common tobacco)  
 C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C;Accession: T04076  
 R;Lermontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.  
 Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997  
 A;Title: Cloning and characterization of a plastidial and a mitochondrial isoform of tobacco  
 A;Reference number: Z15186; MUID:97385200; PMID:9238074  
 A;Accession: T04076  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-504 <LER>  
 A;Cross-references: UNIPROT:O24164; EMBL:Y13466; NID:g2370334; PIDN:CAA73866.1; PID:g237  
 A;Experimental source: strain SRI  
 C;Genetics:  
 A;Gene: PPXII  
 C;Function:  
 A;Description: EC 1.3.3.4 [validated, MUID:97385200]  
 A;Pathway: tetrapyrrole synthesis  
 C;Keywords: mitochondrion; oxidoreductase

Query Match 72.7%; Score 32; DB 2; Length 504;  
 Best Local Similarity 62.5%; Pred. No. 65;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQLMEPV 9

Db 128 LQMLEPI 135

#### RESULT 15

F86251  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C;Accession: F86251  
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: F86251  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-536 <STO>  
 A;Cross-references: UNIPROT:Q9SA99; GB:AE005172; NID:g4835780; PIDN:AAD30246.1; GSPDB:GNCI

A;Map position: 1  
 C;Superfamily: poppy reticuline oxidase

Query Match 72.7%; Score 32; DB 2; Length 536;  
 Best Local Similarity 55.6%; Pred. No. 69;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9

Db 286 FMRLLQPV 294

Search completed: December 30, 2004, 20:52:19  
 Job time : 3.97229 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:30:54 ; Search time 77.1385 Seconds  
(without alignments)  
439.058 Million cell updates/sec

Title: US-10-017-327-2

Perfect score: 1812

Sequence: 1 MASRKEGTGSTATSSSTAG.....FTAQNLGKLFMAQALQEYNN 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	627.5	34.6	337	2 B86242	hypothetical prote
2	526.5	29.1	461	2 T28786	hypothetical prote
3	383	21.1	357	2 T41716	conserved hypothet
4	180.5	10.0	306	2 S6259	26S proteasome reg
5	166.5	9.2	308	2 T43293	multidrug resistan
6	166.5	9.2	308	2 T44427	hypothetical prote
7	162	8.9	312	2 T33344	hypothetical prote
8	152.5	8.4	357	2 T09261	JUN kinase-activat
9	152	8.4	319	2 S44642	hypothetical prote
10	150	8.3	357	2 C86363	Jun activation dom
11	148	8.2	357	2 T52180	constitutive photo
12	143.5	7.9	281	2 E30112	26S proteasome reg
13	142.5	7.9	360	2 T02934	JUN-activation-dom
14	138.5	7.6	358	2 T52042	constitutive photo
15	137.5	7.6	358	2 H96736	hypothetical prote
16	127.5	7.0	321	1 BWSV4	Mov-34 protein - m
17	127	7.0	302	2 T40490	probable 26S prote
18	126	7.0	334	2 S71820	JUN-activation-dom
19	125.5	6.9	293	2 H84823	26S proteasome reg
20	125.5	6.9	321	2 S65491	26S proteasome reg
21	124	6.8	455	2 S67775	hypothetical prote
22	121.5	6.7	294	2 T20338	hypothetical prote
23	120.5	6.7	324	1 T94154	26S proteasome reg
24	120	6.6	368	2 T29320	hypothetical prote
25	110	6.1	562	2 T49904	hypothetical prote
26	110	6.1	1953	2 S63244	BN11 protein - yea
27	109	6.0	362	2 T33096	hypothetical prote
28	108.5	6.0	1165	2 S27809	Grpase-activating
29	108	6.0	707	2 A46302	PTB-associated spl

RESULT 1

B86242

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: B86242

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86242

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-337 <STO>

A;Cross-references: UNIPROT:Q9C5Z2; GB:A8005172; NID:g4874264; PIDN:AAD31329.1; GSPDB:GN

C;Genetics:

A;Map position: 1

Query Match 34.6%; Score 627.5; DB 2; Length 337;

Best Local Similarity 41.7%; Pred. No. 9.3e-38;

Matches 136; Conservative 67; Mismatches 114; Indels 9; Gaps 5;

Qy 31 SGDSAV----KQVQIDGLVLKIIKHVQEGQTEVQVGLLVGLVEDRLITNCFFPPQ 86

Db 13 SKDRAVAPPLRVQIEGLAVLKIIKHCKE--FSTLTVTGQLLDVGLVETVNCFFPPV 70

Qy 87 HTED-DADFDDEVQYQMEMRSLRHVNIDHLHVGWYQSTYYSFVTRALLDSQFSYQHAIE 145

Db 71 RDDDEIEADGANYQLEMRCLERVVNNVTGVYQSTVGLSGYQTVELIETFMVYQENIK 130

Qy 146 ESVVLIYDPIKTAQGSLSLKAYRLTPKLMVECKEKFSPKALKKANITFFYMFPEVPVI 205

Db 131 RCVCIIYDPSKADLGLVALLKALKLSDSFMELYRGNGFTGKLRKFNFSWMDIFEIPIKV 190

Qy 206 KNSHLINVLWELEKKSAAVAD-KHELLSLASSNHLGNLQLLMRDVDEMSQDIVKYNTYM 264

Db 191 SNSALVSFAFMTLSTDTTPVSGDYDRLHSSTTPPLENNMEFLIKCMDLSNEQKQFYVY 250

Qy 265 RNTSKQQQKHQYQRRQENMQRSRGEPLPEEDLSK-LFKPQPQPPARMDSLLIAGQI 323

Db 251 RNLRSQAQQAQWIKQKRTENMARKSAGEPLPEEDPSNPIFKAIPEPSRLSEPLITNOV 310

Qy 324 NTYQCNKKEFTAQNLGKLFMAQALQE 349

Db 311 SNFCGQINGVAGQNFSLRYLTJLALHD 336

transcription acti  
jun activation dom  
26S proteasome reg  
DNA repair protein  
hypothetical prote  
Gunc protein - xan  
R-cadherin precurs  
E1A-associated cyc  
H4 protein - human  
dead ringer nuclea  
protein kinase SK2  
bullous pemphigoid  
methyl-accepting c  
splicing factor ho  
histidine kinase h  
hypothetical prote



Qy 127 SFVTRALDLSQFSYQHAIEESVVLIIYDPIKTAQGSLSLKAYRL--TPKLMVECKEKDFSP 184  
Db 116 CWLSSVDVNTQKSPQLNSRAVAVVVDPIQSVKGVVIDAFRLIDTGLINLNLEPRQTT 175  
Qy 185 EA-LKKANIYFYMFEEVPIVKNSHLINVLWMELEKKSADVADKHELLSLASSNHLGN 242  
Db 176 NTGLLNKANI-----QALLHGLNRHYSLNI-----DYHK-----TAKE 209  
Qy 243 LQLLMDRVDEMSQDIKVKNY-----MRNTSKQQQKHQVQORRQENMQRSRGEP 295  
Db 210 TKMLNHLKEQWQSLKMYDEEESNLATKSMVKIAEQYSKRIEEK-----E 260  
Qy 296 LPEEDLSKLFKPPPPAR 313  
Db 261 LTEBELKTRYVGRQDPKK 278

RESULT 5  
T43293  
multidrug resistance protein skel1 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T43293; T38630  
R:Usui, T.; Yoshida, M.; Honda, A.; Beppu, T.; Horinouchi, S.  
Gene 161, 93-96, 1995  
A:Title: A K-252a-resistance gene, skel1+, encodes a protein similar to the Caenorhabditis  
A:Reference number: 222399; MUID:95369740; PMID:7642144  
A:Accession: T43293  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-308 <USU>  
A:Cross-references: UNIPROT:P41878; EMBL:D45047; NID:G624935; PIDN:BAA08087.1; PID:G6249  
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: 221722  
A:Accession: T38630  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-308 <MCL>  
A:Cross-references: EMBL:Z98979; PIDN:CAB11697.1; GSPDB:GN00066; SPDB:SPAC31G5.13  
A:Experimental source: strain 972h-; cosmid c31G5  
C:Genetics:  
A:Gene: skel1+; SPAC31G5.13  
A:Map position: 1  
C:Function:  
A:Description: essential for vegetative growth; confers multidrug resistance

Query Match 9.2%; Score 166.5; DB 2; Length 308;  
Best Local Similarity 24.4%; Pred. No. 0.00012;  
Matches 68; Conservative 56; Mismatches 116; Indels 39; Gaps 13;

Qy 23 GKGGKGGSGDSAV-----KQVQIDGLVVLKTIKHVQEGQTEV-VQGVILGLVVED-RL 76  
Db 10 GARMGTGMGQPLVDNSECVYISLALLKMLRHR---HGTPMEVGMGLMGEFVDDFTV 66  
Qy 77 EITNCFPPPOH--TEDDADFDEVOYQEMNMRSLRHVNIDHLHVGWYQS--TYYGSFVTRALL 134  
Db 67 RVVDVFAMPQSGTGVSVAEVDPVFKNMMDMLKQTGRPEVMVGVNHPGFCWMLSSVDI 126  
Qy 135 DSQFSYQHAIEESVVLIIYDPIKTAQGSLSLKAYRL--TPKLMVECKEKDFSPAL---KKA 190  
Db 127 NTQGSFQELTPRAVAVVVDPIQSVKGVVIDAFRLINPSTLMWGQEPRTQTTNLSLGHINKP 186  
Qy 191 NIT-----FYMFEVPIVKNSHLINVLWMELEKKSADVADKHELL-----SLASSNHLG 240  
Db 187 SIQALIHGLGRHYSLRINYKTELEELMLNLHKQPAW---HGLLENFNFAAEKKNHAS 243  
Qy 241 KNLQLLMDRVDEMSQDIKVKNYTMNRN--TSKQQQKHQY 277  
Db 244 -----IDKMSLSLSEQYTERVQNEVTLSPQLRIQY 273

RESULT 6  
T33344  
hypothetical protein K07D4.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
R:Henkhaus, J.; Wohldmann, P.  
submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of C. elegans cosmid K07D4.  
A:Reference number: 221327  
A:Accession: T33344  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-312 <HEN>  
A:Cross-references: UNIPROT:O76577; EMBL:AF077534; PIDN:AAC26287.1; GSPDB:GN00020; CRSP:  
A:Experimental source: strain Bristol N2; clone K07D4  
C:Genetics:  
A:Gene: CESP:K07D4.3  
A:Map position: 2  
A:Introns: 16/3; 82/3; 192/3

Query Match 8.9%; Score 162; DB 2; Length 312;  
Best Local Similarity 23.3%; Pred. No. 0.00026;  
Matches 67; Conservative 59; Mismatches 130; Indels 42; Gaps 14;

Qy 20 GAAGKGGKGGSGDSAVKQVQ-----IDGLVVLKTIKHVQEGQTEVQGVILGLVVE 73  
Db 8 GGLGGLNLTGFGANPDQSNQVDTSETVYISSLALLKMLKH-GRAGVPHEVM-GLMLGFEVD 65  
Qy 74 D-RLEITNCFPPPOH--TEDDADFDEVOYQEMNMRSLRHVNIDHLHVGWYQS--TYYGSPVT 130

T44427  
hypothetical protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T44427  
R:Shimanuki, M.; Saka, Y.; Yanagida, M.; Toda, T.  
J. Cell Sci. 108, 569-579, 1995  
A:Title: A novel essential fission yeast gene padl+ positively regulates papl+-dependent  
A:Reference number: 222764; MUID:95286704; PMID:7769002  
A:Accession: T44427  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-308 <SHI>  
A:Cross-references: UNIPROT:P41878; EMBL:D31731; NID:G497632; PIDN:BAA06529.1; PID:G4976  
C:Genetics:  
A:Gene: padl+

Query Match 9.2%; Score 166.5; DB 2; Length 308;  
Best Local Similarity 24.4%; Pred. No. 0.00012;  
Matches 68; Conservative 56; Mismatches 116; Indels 39; Gaps 13;

Qy 23 GKGGKGGSGDSAV-----KQVQIDGLVVLKTIKHVQEGQTEV-VQGVILGLVVED-RL 76  
Db 10 GARMGTGMGQPLVDNSECVYISLALLKMLRHR---HGTPMEVGMGLMGEFVDDFTV 66  
Qy 77 EITNCFPPPOH--TEDDADFDEVOYQEMNMRSLRHVNIDHLHVGWYQS--TYYGSFVTRALL 134  
Db 67 RVVDVFAMPQSGTGVSVAEVDPVFKNMMDMLKQTGRPEVMVGVNHPGFCWMLSSVDI 126  
Qy 135 DSQFSYQHAIEESVVLIIYDPIKTAQGSLSLKAYRL--TPKLMVECKEKDFSPAL---KKA 190  
Db 127 NTQGSFQELTPRAVAVVVDPIQSVKGVVIDAFRLINPSTLMWGQEPRTQTTNLSLGHINKP 186  
Qy 191 NIT-----FYMFEVPIVKNSHLINVLWMELEKKSADVADKHELL-----SLASSNHLG 240  
Db 187 SIQALIHGLGRHYSLRINYKTELEELMLNLHKQPAW---HGLLENFNFAAEKKNHAS 243  
Qy 241 KNLQLLMDRVDEMSQDIKVKNYTMNRN--TSKQQQKHQY 277  
Db 244 -----IDKMSLSLSEQYTERVQNEVTLSPQLRIQY 273

RESULT 7  
T33344  
hypothetical protein K07D4.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
R:Henkhaus, J.; Wohldmann, P.  
submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of C. elegans cosmid K07D4.  
A:Reference number: 221327  
A:Accession: T33344  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-312 <HEN>  
A:Cross-references: UNIPROT:O76577; EMBL:AF077534; PIDN:AAC26287.1; GSPDB:GN00020; CRSP:  
A:Experimental source: strain Bristol N2; clone K07D4  
C:Genetics:  
A:Gene: CESP:K07D4.3  
A:Map position: 2  
A:Introns: 16/3; 82/3; 192/3

Query Match 8.9%; Score 162; DB 2; Length 312;  
Best Local Similarity 23.3%; Pred. No. 0.00026;  
Matches 67; Conservative 59; Mismatches 130; Indels 42; Gaps 14;

Qy 20 GAAGKGGKGGSGDSAVKQVQ-----IDGLVVLKTIKHVQEGQTEVQGVILGLVVE 73  
Db 8 GGLGGLNLTGFGANPDQSNQVDTSETVYISSLALLKMLKH-GRAGVPHEVM-GLMLGFEVD 65  
Qy 74 D-RLEITNCFPPPOH--TEDDADFDEVOYQEMNMRSLRHVNIDHLHVGWYQS--TYYGSPVT 130

Db 66 DYTNNVDVFAMPQSGTGVSAEADVPFQAKMLDMLKOTGRPEMVVGVYHSHPGGCWLS 125  
Qy 131 RALLDSQSYQHAITEESVLLIYDPIKTAQGSLSLKAYR-LTPKLMVEVCKEKDFSE---A 186  
Db 126 GVDINTQSQSFALSDRAVAVVVDPIQSVKGVKVIDAFRTINPQSMALNQEPQRTTSLGH 185  
Qy 187 LKKNIT-----FYMPEEPIVINKSHL-----INVLMW-----ELEKKSADADKH 228  
Db 186 LQKPSIQALHGLNRHYYSIIPIARTHLEQKMLLNKLKUSMDAVSVENYSKCGEQNKE 245  
Qy 229 ELLSLASSNHGKNL-QLLMDRVDSEMSQDIVKYNTYMYRNTSKQQQKX 275  
Db 246 HLKAMLK---LAKNYKALEKXMTDQELA-----IKNVGKMPKXH 285

RESULT 8  
T09261  
JUN kinase-activation-domain-binding protein homolog - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09261  
R:Rajasekhar, V.K.; Reddy, J.T.; Gonzales, R.A.; Dixon, R.A.  
submitted to the EMBL Data Library, July 1997  
A:Description: Putative plant JUN activation domain binding protein.  
A:Reference number: Z16633  
A:Accession: T09261  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-357 <RAJ>  
A:Cross-references: UNIPROT:O81090; EMBL:AF014375; NID:g3320378; PID:g3320379  
C:Superfamily: Caenorhabditis elegans hypothetical protein B0547.1

Query Match 8.4%; Score 152.5; DB 2; Length 357;  
Best Local Similarity 21.1%; Pred. No. 0.0015;  
Matches 71; Conservative 67; Mismatches 133; Indels 65; Gaps 12;  
Qy 19 AGAAGKGGKGGSD-SAVKQVQIDGLVLIKIHYQEGGQTEVVGQVLLGLVWEDRL 77  
Db 39 AGQAEFQDKPWPANDPHYKRVKISALALLKVVH--ARSGGIIEVWGLMGKTDADSII 96  
Qy 78 ITNCFPPF---QHTEDDADFEVQYQEMMRSRLHVNIDHLHVGWYQS-TYIGSFVTRAL 133  
Db 97 VMDAFALPVEGTETRVNAQADAVEYMDVYSQTNKQAGRLNVVGVYHSHPGYGCWLSGID 156  
Qy 134 LDSQSYQHAITEESVLLIYDPIKT-AQGSLSLKAYRLTP-----KLMEVCKE 179  
Db 157 VSTQMLNQEQEPFLAVVIDPTFVTSAGKVEIGAFRTYPEGYKPADDPISYQTIPINKI 216  
Qy 180 KDFSPEALKKANITFEYMFEEVPIVINKSHLINVLMMELEKKSADADKHLSLASSNHL 239  
Db 217 EDFGVHCKQYSLDITYFKSSL-----DSHLDD-LLW-----NKYVWNTLSSSPLL 261  
Qy 240 GKNLQLLMDRVDSEMSQDIVKYNTYMYRNTSKQQQKXHQYQRRQENMQRSRGP----- 294  
Db 262 G-NGDIVAGQISDLAKEL-----EQANQLAHSRFGPLVAPT 297  
Qy 295 PLPEEDLSKLFKPPQPPARMDSLLIAGQINTYQNI 330  
Db 298 PRKKEESPLAKITRDSAKITVEQVHGLMSQVIKDI 333

RESULT 9  
S44642  
hypothetical protein F37A4.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S44642; T16290  
R:Fulton, L.  
submitted to the EMBL Data Library, February 1994  
A:Description: The sequence of C. elegans cosmid F37A4.  
A:Reference number: S44638  
A:Accession: S44642  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-319 <FUL>  
A:Cross-references: UNIPROT:P41883; EMBL:U00032; NID:g458958; PID:g458963; PIDN:AAA50633  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CBSP:P37A4.5  
A:Introns: 28/2; 109/2; 295/3

Query Match 8.4%; Score 152; DB 2; Length 319;  
Best Local Similarity 24.2%; Pred. No. 0.0014;  
Matches 59; Conservative 55; Mismatches 96; Indels 34; Gaps 10;  
Qy 39 VQIDGLVLIKIHYQEGGQTEVVGQVLLGLVVED-RLEINTNCFPPQH-TEDDADFE 96  
Db 29 VNISSLALLKMLRH-ARSGIPLVM-GLMLGDFVDYTTINTVDVFAMPQSGTSVTVESVD 86  
Qy 97 VOYQEMMRSRLHVNIDHLHVGWYQS-TYIGSFVTRALLDSQSYQHAITEESVLLIYDPI 155  
Db 87 FVYQTKHMDLLKLVGTENVVGVYHSHPGFGLWSVDVNTQSFALHPRAVAVVVDPI 146  
Qy 156 KTAQGSLSLKAYR-LTPKLMVEVCKEKDFSPALKKANITFEYMFEEVPIVINKSHLINV 214  
Db 147 QSVKGVKWLDAFRSVNPLNLIQI---RPLAPTAEPRQTT-----N 183  
Qy 215 MWELEKKSADADKHEI-----LSLASSNHGKNLQLLMDRVDSEMS-QDIVKYNTYMYRNTSK 269  
Db 184 LGHLTKPSLISVVHGLGTKYSLNAVYRMSNEQKMLCMLKKSWYDQLNMSTYSELEKK 243  
Qy 270 QQQQ 273  
Db 244 QEEK 247

RESULT 10  
C86363

JUN activation domain binding protein homolog - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: C86363  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anan, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroso, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C86363  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-357 <STO>  
A:Cross-references: UNIPROT:O23130; GB:AE005172; NID:g2462824; PIDN:AAB72159.1; GSPDB:GN1  
C:Genetics:  
A:Map position: 1  
C:Superfamily: Caenorhabditis elegans hypothetical protein B0547.1

Query Match 8.3%; Score 150; DB 2; Length 357;  
Best Local Similarity 21.1%; Pred. No. 0.0023;  
Matches 56; Conservative 62; Mismatches 112; Indels 36; Gaps 9;  
Qy 37 KOVQIDGLVLIKIHYQEGGQTEVVGQVLLGLVVEDRLRLEITNCFPPF---QHTEDDAD 93  
Db 57 KRVLISALLKMLVVH--ARSGGIIEVWGLMGKTEGTIIIVMDAFALPVEGTETRVNAQ 114  
Qy 94 FDEVQYQEMMRSRLHVNIDHLHVGWYQS-TYIGSFVTRALLDSQSYQHAITEESVLLIY 152  
Db 115 SDAYEVVVEYQSTKLGRLENVGVYHSHPGYGCWLSGIDVSTQMLNQEQEPFLAVVI 174  
Qy 153 DPIKT-AQGSLSLKAYRLTP-----KLMEVCKEKDFSPALKKANITFEYMF 198

Db 175 DPTRTVSAGKVEIGAFRTYPBGHKISDDHVSVEYQTIPLNKIEDFGVHCKQYYSLDITYFK 234

Qy 199 EEVPIVKNSHLINVLWMELEKKSAAVDKHELLSLASSNHLGKNLQLLMDRVDMSQDIV 258

Db 235 SSL-----DSHLLD-LLW-----NKYWNTLSSSPLLG-NGDYVAGQISDLAEKLE 278

Qy 259 KNTYMRNTSKQOQKHQYQORROOE 284

Db 279 QAESQLANSRYGGIAPAGHORRKEDE 304

RESULT 11

T52180

constitutive photomorphogenic 9 complex chain AJH1 [validated] - Arabidopsis thaliana

N;Alternate names: COP9 complex chain AJH1; transcription regulator AJH1

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C;Accession: T52180

R;Kwok, S.F.; Solano, R.; Taube, T.; Chamovitz, D.A.; Matsui, M.; Ecker, J.R.; Deng, X.W

Plant Cell 10, 1779-1790, 1998

A;Title: Arabidopsis homologs of a c-Jun coactivator are present both in monomeric form

A;Reference number: Z25918; MUID:99030187; PMID:9811788

A;Accession: T52180

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-357 <KWO>

A;Cross-references: UNIPROT:O82524; EMBL:AF087413; PIDN:AAC36344.1

C;Complex: octamer; consisting of COP9, FUS6, AJH2/1 and others; AJH1 also exists in a m

C;Function:

A;Description: involved in repression of light-mediated development; may function in mod

C;Superfamily: Caenorhabditis elegans hypothetical protein B0547.1

Query Match 8.2%; Score 148; DB 2; Length 357;

Best Local Similarity 20.7%; Pred. No. 0.0032;

Matches 55; Conservative 62; Mismatches 113; Indels 36; Gaps 8;

Qy 37 KQVQIDGLVLKIIKHQESQGTETVQVGLLGLVDEDRLEITNCFFPP---QHTEDDAD 93

Db 57 KRVTISALLKMWVH--ARSGGTIEIMGLMQGTEGDTIIVMDAFALPVEGTETRVNAQ 114

Qy 94 FDEVOYQMEMRSLRHVNIDHLHVGWYQS--TYTSGSFVTRALLDSQFSYQHAIESVVLII 152

Db 115 SDAYEYVVEYSQTKLAGRLNVVGVYHSHPGVGCWLSGIDVSTQMLNQOQEPFLAVI 174

Qy 153 DPIKT-AQGSLSLKAYRLTP-----KLMEVCKEKDPSPEALKKANITFEYMF 198

Db 175 DPTRTVSAGKVEIGAFRTYPBGHKISDDHVSVEYQTIPLNKIEDFGVHCKQYYSLDITYFK 234

Qy 199 EEVPIVKNSHLINVLWMELEKKSAAVDKHELLSLASSNHLGKNLQLLMDRVDMSQDIV 258

Db 235 SSL-----DSHLLD-LLW-----GNKYWNTLSSSPLLG-NGDYVAGQISDLAEKLE 278

Qy 259 KNTYMRNTSKQOQKHQYQORROOE 284

Db 279 QAESQLANSRYGGIAPAGHORRKEDE 304

RESULT 12

E90112

26S proteasome regulatory subunit [imported] - Guillardia theta nucleomorph

C;Species: nucleomorph Guillardia theta

A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C;Accession: E90112

R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei

Nature 410, 1091-1096, 2001

A;Title: The highly reduced genome of an enslaved algal nucleus.

A;Reference number: A99082; MUID:11323671; PMID:11323671

A;Accession: E90112

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-281 <DOU>

A;Cross-references: UNIPROT:Q9AW06; GB:AJ010592; NID:g12580747; PIDN:CAC27065.1; GSPDB:G

C;Genetics:

A;Map position: 2

A;Genome: nucleomorph

C;Keywords: nucleomorph

Query Match 7.9%; Score 143.5; DB 2; Length 281;

Best Local Similarity 26.4%; Pred. No. 0.0049;

Matches 39; Conservative 43; Mismatches 61; Indels 5; Gaps 5;

Qy 37 KQVQIDGLVLKIIKHQESQGTETVQVGLLGLVDEDRLEITNCFFPPQ-HTEDDADF 94

Db 15 ESINISSIALIKMLKH-SKAGVPVEVM-GLLGNFVDEINISVNDVAMPOTGTGISVES 72

Qy 95 DEVOYQMEMRSLRHVNIDHLHVGWYQS--TYTSGSFVTRALLDSQFSYQHAIESVVLII 153

Db 73 LDPSFQTKMLDLSQLGNKSLIVGVYHSHPGVGCWLSGIDVSTQMLNQOQEPFLAVI 132

Qy 154 PIKTAQGSLSLKAYRLTPKLMVEVCKEKD 181

Db 133 PIQSFKNRMFIEAYRLPSGKNQKVLERE 160

RESULT 13

T02934

JUN-activation-domain-binding protein homolog - rice

C;Species: Oryza sativa (rice)

C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C;Accession: T02934

R;Meijer, A.H.; de Kam, R.J.; Qin, L.; Hoge, J.H.C.

submitted to the EMBL Data Library, June 1998

A;Reference number: Z14779

A;Accession: T02934

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-360 <MEI>

A;Cross-references: UNIPROT:Q8H936; UNIPROT:O81388; EMBL:AF072849; NID:g3420298; PID:g934

A;Experimental source: subsp. Indica

C;Genetics:

A;Gene: jab1

C;Superfamily: Caenorhabditis elegans hypothetical protein B0547.1

Query Match 7.9%; Score 142.5; DB 2; Length 360;

Best Local Similarity 21.6%; Pred. No. 0.0081;

Matches 57; Conservative 55; Mismatches 109; Indels 43; Gaps 10;

Qy 37 KQVQIDGLVLKIIKHQESQGTETVQVGLLGLVDEDRLEITNCFFPP---QHTEDDAD 93

Db 58 RRAKISALLKMWVH--ARAGGTIEVGLMQGCEGDAIWMDAFALPVEGTETRVNAQ 115

Qy 94 FDEVOYQMEMRSLRHVNIDHLHVGWYQS--TYTSGSFVTRALLDSQFSYQHAIESVVLII 152

Db 116 ADAYEYVVEYSTNKQAGRLNVVGVYHSHPGVGCWLSGIDVSTQMLNQOQEPFLAVI 175

Qy 153 DPIKT-AQGSLSLKAYRLTPK-----LMEVCKEKDPSPEALKKANITFEYMF 198

Db 176 DPTRTVSAGKVEIGAFRTYPKDYKPPDPVSEYQTIPLNKIEDFGVHCKQYVALDITYFK 235

Qy 199 EEVPIVKNSHLINVLWMELEKKSAAVDKHELLSLASSNHLGKNLQL-----LMDRVD 252

Db 236 SSL-----DSHLLD-LLW-----NKYWNTLSSSPLLGNGRDYVAGQIFDLADKLEQ 280

Qy 253 MSQDIV--KNTYMRNTSKOQOQK 274

Db 281 AEGQLANSRYGMLMPSQRKKEQE 304

RESULT 14

T52042

constitutive photomorphogenic 9 complex chain AJH2 [validated] - Arabidopsis thaliana

N;Alternate names: COP9 complex chain AJH2; transcription regulator AJH2

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C/Accession: T52042  
R/Kwok, S.F.; Solano, R.; Tsuge, T.; Chamovitz, D.A.; Matsui, M.; Ecker, J.R.; Deng, X.W.  
Plant Cell 10, 1779-1790, 1998  
A/Title: Arabidopsis homologs of a c-Jun coactivator are present both in monomeric form  
s.  
A/Reference number: Z25918; MUID:99030187; PMID:9811788  
A/Accession: T52042  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: mRNA  
A/Residues: 1-358 <KWO>  
A/Cross-references: UNIPROT:O82523; EMBL:AF087412; PIDD:AAC36343.1  
C/Complex: octamer; consisting of COP9, FUS6, AJH2/1 and others; AJH2 also exists in a "C/Function:  
A/Description: involved in repression of light-mediated development; may function in mod  
C/Superfamily: Caenorhabditis elegans hypothetical protein B0547.1

Query Match	7.6%; Score 138.5; DB 2; Length 358;
Best Local Similarity	21.3%; Pred. No. 0.016;
Matches	72; Conservative 70; Mismatches 123; Indels 73; Gaps 15;
Qy	37 KQVQIDGLAVLKIHKHQEBSQGTVEVQGVLLGLWVEDRLIETNCFFPP---QHTEDDAD 93
Db	57 KRKVISALALLKMWVH---ARSGGTIEIEMGLMGKGTGDTIIVMDAFALPVEGTETRVNAQ 114
Qy	94 FDEYQYQEMMRSLRHVNIDHLHVGWYQS--TYYGSPVTRALLDSQFSYQHAIESVVLV 152
Db	115 DDAVEYVVEYSQTNKLAGRLNVGVGTHSPHGVCWLSGIDVSTQTNLNQHQRPFLAVI 174
Qy	153 DPIKT-AQGSLSLKAYRLTPK-----LMEVCKEKDQSPALKKANITFEYMF 198
Db	175 DPTRTTSAGKVEIGAFRTYSKGYPDPPEVSEYQTIPLNKIEDFGVHCKQYVSLDVTYPK 234
Qy	199 EEPVPIVKNSHLNLVLMWELEKKSAVADKHELLSLASSNHGLKNLQLLMDRVDEMSSDIV 258
Db	235 SSL-----DSHLLD-LLW-----NKYVWNTLSSPLLG-NGDYVAGQISDLAEKL- 277
Qy	259 KYNTYMRNTSKQOQKHQYQORRQENMQRSRGERP-----LPEEDLSKLFPKPPQPPARM 314
Db	278 -----EQAESHLVQSR-----FGGVPPSLHKKKEDESQUTKITRDSAKI 317
Qy	315 DSLLIAGQINTYCNKTEFTAQNLGLFMAQALQEVNN 352
Db	318 TVEQVTHGLMS---QVIKD-----ELF---NSMROSN 343

RESULT 15

H96736

hypothetical protein F3117.12 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: H96736  
R:Thelodopsis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizlar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; PMID:21016719; PMID:11130712

Query Match  
7.6%; Score 137.5; DB 2; Length 358;

Best Local Similarity	21.3%;	Pred. No. 0.018;	
Matches	72;	Conservative	70;
		Mismatches	123;
		Indels	73;
		Gaps	15;
Qy	37	KQVQIDGLVLIKIKHYQEBEGQCTEVVQVGLLGLVVEDRL EITNCFFP	--QHTEDDAD 93
Db	57	KRVKISALALLKVVVH--ARSGGTIEIMGLMQGKTGDTIIVMDALP	VEGTETRVNAQ 114
Qy	94	FDEVQIQMEMRSLRHVNIDHLHVGVQYS--TYYGSPVTRALLDSQ	SYQHAIEESVVLII 152
Db	115	DDAYEYVMEYSQTNKAGRLNVGVYGHSPGYCGCLSGIDVSTQR	LNOHQHQPFFLAVVI 174
Qy	153	DP1KTT-AQGSLSLKAURLTPK-----LMEVCCKDFSP	EALKKANITEYMF 198
Db	175	DPRTVTSAGKVEIGARTSYKGYKPPDEPSEVQTIPLNKIEDFGV	CHQKYQYSLDVTYFK 234
Qy	199	EEVPIVINKSHLINVLWBLEKKSAVADKHELLSLASSNHLGKNL	QLLMDRVDEMSSQDIV 258
Db	235	SSL-----DSHLLD-LLW-----NKYWNLTSSSSPLL	G-NGDYVAGQISDLAEKL- 277
Qy	259	KYNTYMRNTSKQOQKHQYQORRQEQNMQRQSRGEPP---	LPEEDSLKLFKPPQPPARM 314
Db	278	-----EQAESHLVQSR-----PGGVVPPSLHKHKKED	ESQTKITRDSAKI 317
Qy	315	DSLLIAGQINTVQCNTKEFTQAQNLGKLFMAQALQEYNN	352
Db	318	TVEQVHGLMS--QVTKD-----ELF--NSMROSNN	343

Search completed: December 30, 2004, 20:52:17  
Job time : 79.1385 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:29:59 ; Search time 413.179 Seconds  
(without alignments)  
490.180 Million cell updates/sec

Title: US-10-017-327-2

Perfect score: 1812

Sequence: 1 MASRKEGTGSTATSSSTAG.....FTAQNLGKLFMAQALOBYNN 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1812	100.0	352	1 IF33 HUMAN	O15372 homo sapien
2	1812	100.0	352	2 CAG33187	Cag33187 homo sapi
3	1772	97.8	352	2 Q6P9U8	Q6P9U8 rattus norv
4	1772	97.8	352	2 AAH60586	AAH60586 rattus no
5	1762	97.2	352	1 IF33 MOUSE	Q1wk2 mus musculu
6	1759	97.1	352	2 Q8BTX5	Q8BTX5 mus musculu
7	1463.5	80.8	335	2 Q6P381	Q6P381 xenopus tro
8	1463.5	80.8	335	2 AAH64151	AAH64151 xenopus t
9	848	46.8	337	2 Q9GV27	Q9GV27 bombyx mori
10	830.5	45.8	337	2 Q7PVR3	Q7PVR3 anopheles g
11	771	42.5	337	2 Q7KLY2	Q7KLY2 drosophila
12	771	42.5	337	2 AAD46836	AAD46836 drosophil
13	771	42.5	338	2 Q9U904	Q9U904 drosophila
14	771	42.5	338	2 AAF52210	AAF52210 drosophil
15	697	38.5	149	2 Q6Y126	Q6Y126 pagrus majo
16	697	38.5	149	2 AAP20218	AAP20218 pagrus ma
17	666.5	36.8	347	2 Q7XMG6	Q7XMG6 oryza sativ
18	627.5	34.6	337	1 IF33 ARATH	Q9C522 arabidopsis
19	627.5	34.6	337	2 AAN31904	AAN31904 arabidops
20	627.5	34.6	337	2 AAM64888	AAM64888 arabidops
21	566	31.2	365	2 O01974	O01974 caenorhabdi
22	508	28.0	359	2 Q7S9Y9	Q7S9Y9 neurospora
23	487.5	26.9	347	2 Q6C113	Q6C113 yarrowia li
24	383	21.1	357	2 Q9UT48	Q9UT48 schizosacch
25	348	19.2	335	2 Q6BMB6	Q6BMB6 debaryomyce
26	184.5	10.2	311	2 Q6BK38	Q6BK38 debaryomyce
27	184	10.2	70	2 Q8BMZ8	Q8BMZ8 mus musculu
28	183	10.1	310	2 Q00487	Q00487 homo sapien
29	183	10.1	310	2 Q9CZV6	Q9CZV6 m mus muscu
30	183	10.1	310	2 AAH66336	AAH66336 homo sapi
31	180.5	10.0	306	1 MPRI_YEAST	P43588 saccharomyc

## RESULT 1

IF33 HUMAN  
ID IF33 HUMAN STANDARD; PRT; 352 AA.  
AC O15372;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Eukaryotic translation initiation factor 3 subunit 3 (eIF-3 gamma)  
DE (eIF3 p40 subunit) (eIF3h).  
GN Name=EIF3S3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=98001678; PubMed=9341143;  
RA Asano K., Vornlocher H.-P., Richter-Cook N.J., Merrick W.C.,  
RA Hinnebusch A.G., Hershey J.W.B.;  
RT "Structure of cDNAs encoding human eukaryotic initiation factor 3  
subunits. Possible roles in RNA binding and macromolecular assembly.";  
RL J. Biol. Chem. 272:27042-27052(1997).  
[2]  
SEQUENCE FROM N.A.  
RA Schmidt O.G., von Holtum D., Gross S., Horsthemke B., Lueddecke H.-J.;  
RT "The gene encoding the p40 subunit of the translation initiation  
factor eif3 has 8 exons, maps to the Langer-Giedion syndrome region on  
chromosome 8q24, but is not the TRPS gene.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Richmond J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Q6C0A6 yarrowia li  
Q7x92 xenopus lae  
Q750E9 ashbya goss  
AAS54495 ashbya go  
Q7Q8I2 anopheles g  
Q9V3H2 drosophila  
Q8J0T5 saccharomyc  
Q35593 mus musculu  
O16154 schistosoma  
Q8QY3 encaphalito  
Q6Eks1 candida gla  
Q7R3D6 giardia lam  
Q966Y7 aphrocallis  
Q9N9W7 giardia lam





```
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060586; AAH60586.1; -.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000555; Mov34_MPN_PAD1.
DR Pfam; PF01398; Mov34; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Initiation factor.
SQ SEQUENCE 352 AA; 39905 MW; C06307269ADB343 CRC64;
Query Match 97.8%; Score 1772; DB 2; Length 352;
Best Local Similarity 97.7%; Pred. No. 2.1e-116;
Matches 344; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MASKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHYYEGGGT 60
Db 1 MASKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHYYEGGGT 60
Qy 61 EVQGVLLGLVVEVDRLEITNCPFPQHTEDDADFEVQYQEMMRSLRHVNIDHLHVGWY 120
Db 61 EVQGVLLGLVVEVDRLEITNCPFPQHTEDDADFEVQYQEMMRSLRHVNIDHLHVGWY 120
Qy 121 QSTYGSFVTRALLDSQFSYQHAIEESVVLIDYPIKTAQGSLSLKAYRLTPKLMVECKEK 180
Db 121 QSTYGSFVTRALLDSQFSYQHAIEESVVLIDYPIKTAQGSLSLKAYRLTPKLMVECKEK 180
Qy 181 DFSPEALKKANITFEHMFEEVPIVKNSHLINVLWLEKEKSAVADKHELLSLASSNHLG 240
Db 181 DFSPEALKKANITFEHMFEEVPIVKNSHLINVLWLEKEKSAVADKHELLSLASSNHLG 240
Qy 241 KNQLLMDRVDEMDSQDIKYNTYMRNSKQQQKHQYQORRQENMQRSRGEPLPEED 300
Db 241 KNQLLMDRVDEMDSQDIKYNTYMRNSKQQQKHQYQORRQENMQRSRGEPLPEED 300
Qy 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQEYNN 352
Db 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQEYNN 352
RESULT 4
AAH60586 PRELIMINARY; PRT; 352 AA.
AC AAH60586;
DT 14-APR-2004 (TRENBLrel. 27, Created)
DT 14-APR-2004 (TRENBLrel. 27, Last sequence update)
DT 14-APR-2004 (TRENBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 3 gamma,
DE 40kDa.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060586; AAH60586.1; -.
KW Initiation factor.
SQ SEQUENCE 352 AA; 39905 MW; C06307269ADB343 CRC64;
Query Match 97.8%; Score 1772; DB 2; Length 352;
Best Local Similarity 97.7%; Pred. No. 2.1e-116;
Matches 344; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MASKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHYYEGGGT 60
Db 1 MASKEGTGSTATSSSTAGAGKGGKGGSDSAVKQVQIDGLVVLKIIKHYYEGGGT 60
Qy 61 EVQGVLLGLVVEVDRLEITNCPFPQHTEDDADFEVQYQEMMRSLRHVNIDHLHVGWY 120
Db 61 EVQGVLLGLVVEVDRLEITNCPFPQHTEDDADFEVQYQEMMRSLRHVNIDHLHVGWY 120
Qy 121 QSTYGSFVTRALLDSQFSYQHAIEESVVLIDYPIKTAQGSLSLKAYRLTPKLMVECKEK 180
Db 121 QSTYGSFVTRALLDSQFSYQHAIEESVVLIDYPIKTAQGSLSLKAYRLTPKLMVECKEK 180
Qy 181 DFSPEALKKANITFEHMFEEVPIVKNSHLINVLWLEKEKSAVADKHELLSLASSNHLG 240
Db 181 DFSPEALKKANITFEHMFEEVPIVKNSHLINVLWLEKEKSAVADKHELLSLASSNHLG 240
Qy 241 KNQLLMDRVDEMDSQDIKYNTYMRNSKQQQKHQYQORRQENMQRSRGEPLPEED 300
Db 241 KNQLLMDRVDEMDSQDIKYNTYMRNSKQQQKHQYQORRQENMQRSRGEPLPEED 300
Qy 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQEYNN 352
Db 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQEYNN 352
RESULT 5
IF33 MOUSE
ID _IF33 MOUSE STANDARD; PRT; 352 AA.
AC Q91WK2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 3 (eIF-3 gamma)
DE (eIF3 p40 subunit) (eIF3h).
GN Name=Eif3e3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye, and Retina;
```

```

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenken C.M., Schuler G.D.,
RA Altachul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,
RA Butterfield A.S., Krywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of
CC methionyl-tRNAi and mRNA. Associates with the p170 subunit of
CC ERF3.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits.
CC -!- SIMILARITY: Contains 1 MPN (JAB/Mov34) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; BC014755; A014755.1; -.
DR MEROPS; M67.971; -.
DR MGD; MGI:1915385; Eif3b3.
DR InterPro; IPR003639; Mov34-1.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD363422; Mov34_1; 1.
DR SMART; SM00232; JAB MPN; 1.
KW Initiation factor; Protein biosynthesis.
SQ SEQUENCE 352 AA; 39832 MW; 96F5AB8E2F41F838 CRC64;

Query Match          97.28; Score 1762; DB 1; Length 352;
Best Local Similarity 97.28; Pred. No. 1.le-115;
Matches 342; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MASRKEGTGSTATSSSSSTAGAGKGGKGGSGSAVKQVQIDGLVWLKIKHYQEGGQT 60
Db 1 MASRKEGTGSTATSSGAGGAGKGGKGGSGSAVKQVQIDGLVWLKIKHYQEGGQT 60
Qy 61 EVVQGVLLGLVDRLEITNCFFPPQHTEDADDFEYQYQEMWRSIRHVNIDHLHVGY 120
Db 61 EVVQGVLLGLVDRLEITNCFFPPQHTEDADDFEYQYQEMWRSIRHVNIDHLHVGY 120
Qy 121 QSTYVYGSFVTRALLDSQFSYQHAIESVLLYPIKTAQGSLSKAYRLTPKLMVEYCKEK 180
Db 121 QSTYVYGSFVTRALLDSQFSYQHAIESVLLYPIKTAQGSLSKAYRLTPKLMVEYCKEK 180
Qy 181 DFPSEALKKANITFEYNFEEVPIVINKSHLINVLMWELEKKSAAVADKHELLSLASSNHILG 240
Db 181 DFPSEALKKANITFEYNFEEVPIVINKSHLINVLMWELEKKSAAVADKHELLSLASSNHILG 240
Qy 241 KNLQLMDRVDENSQDIKYNTYWRNTSKQQQKHQYQQRQENMQRSGRGPPLPEED 300
Db 241 KNLQLMDRVDENSQDIKYNTYWRNTSKQQQKHQYQQRQENMQRSGRGPPLPEED 300
Qy 301 LSKLFFKPPPPARMDSILLIAGQINTYCNQIKFTAGNLGKLFMAQALQEQYNN 352
Db 301 LSKLFFKPPPPARMDSILLIAGQINTYCNQIKFTAGNLGKLFMAQALQEQYNN 352

```

```

RESULT 6
Q8BCTX5
ID Q8BCTX5 PRELIMINARY; PRT; 352 AA.
AC Q8BCTX5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
DE length enriched library, clone:E430016K01 product:eukaryotic
DE translation initiation factor 3, subunit 3 (gamma, 40kD), full insert
DE sequence.
GN Name:Eif3b3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA The RIKEN Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumi T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,

```

RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayaehizaki Y.,  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK088436; BAC40351.1; -;  
DR MGI; 1915385; Elf363.  
DR GO; GO:0003743; F:translation initiation factor activity; IEA.  
DR InterPro; IPR005555; Mov34\_MPN\_PAD1.  
DR Pfam; PF01398; Mov34; 1.  
DR SMART; SM00232; JAB\_MPN; 1.  
KW Initiation factor.  
SQ SEQUENCE 352 AA; 39846 MW; 8GF5A89E2F41F97D CRC64;  
  
Query Match 97.1%; Score 1759; DB 2; Length 352;  
Best Local Similarity 96.9%; Pred. No. 1.8e-115;  
Matches 341; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 1 MASRKEGTGSTATSSSTAGAGKGGKGGSGDSAVKQVQIDGLVVLKIIKHYYEEGGGT 60  
Db 1 MASRKEGTGSTATSSGAGGAGVGKGGKGGSGDSAVKQVQIDGLVVLKIIKHYYEEGGGT 60  
  
Qy 61 EVQGVLLGLVVDRLITNCFFPPQHTEDDADDFEVOYQYEMMRSLRHVNIDHLHVGWY 120  
Db 61 EVQGVLLGLVVDRLITNCFFPPQHTEDDADDFEVOYQYEMMRSLRHVNIDHLHVGWY 120  
  
Qy 121 QSTYVGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVECKEK 180  
Db 121 QSTYVGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVECKEK 180  
  
Qy 181 DFPSEALKKANITFEYMFEEVPIVKNKSHLINVLMWELEKKSASAVADKHELLSLASSNHLG 240  
Db 181 DFPSEALKKASITFEHMFEEVPIVKNKSHLINVLMWELEKKSASAVADKHELLSLASSNHLG 240  
  
Qy 241 KNLQLLMDRVDEMDSQDIYKNTYMRNTSKQOQKHQYQORQOENMQORSGEPPLPEED 300  
Db 241 KSLQLLMDRVDEMDSQDIYKNTYMRNTSKQOQKHQYQORQOENMQORSGEPPLPEED 300  
  
Qy 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQEYNN 352  
Db 301 LSKLFPKHPQAPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQEYNN 352  
  
RESULT 7  
Q6P381 ID Q6P381 PRELIMINARY; PRT; 335 AA.  
AC Q6P381  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein MGC75580.  
GN Name=MGC75580;  
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Schaefer C.F., Bhat N.K.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC064151; AAH64151.1; -;  
DR InterPro; IPR005555; Mov34\_MPN\_PAD1.  
DR Pfam; PF01398; Mov34; 1.  
DR SMART; SM00232; JAB\_MPN; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 335 AA; 38538 MW; 18D4D446D6561BE9 CRC64;  
  
Query Match 80.8%; Score 1463.5; DB 2; Length 335;  
Best Local Similarity 80.8%; Pred. No. 9.8e-95;  
Matches 282; Conservative 29; Mismatches 21; Indels 17; Gaps 1;  
  
Qy 1 MASRKEGTGSTATSSSTAGAGKGGKGGSGDSAVKQVQIDGLVVLKIIKHYYEEGGGT 60  
Db 1 MAARTREGTATVA-----ETAVKQVQIDGMVVLKIIKHYYEEGGHGS 43  
  
Qy 61 EVQGVLLGLVVDRLITNCFFPPQHTEDDADDFEVOYQYEMMRSLRHVNIDHLHVGWY 120  
Db 44 EVQGVLLGLVVDRLITNCFFPPQHTEDDADDFEVOYQYEMMRSLRHVNIDHLHVGWY 103  
  
Qy 121 QSTYVGSFVTRALLDSQFSYQHAIEESVLLIYDPIKTAQGSLSLKAYRLTPKLMVECKEK 180  
Db 104 QSTYVGFVSRALLDSQFSYQHAIEESVLLIYDPIKTCGGSLSLKAYRLTPKLMVECKEK 163  
  
Qy 181 DFPSEALKKANITFEYMFEEVPIVKNKSHLINVLMWELEKKSASAVADKHELLSLASSNHLG 240  
Db 164 DFPSEAGLKKANIAPEDFEVPFIIKNKSHLINVLAELDKKAPVTEKHELLNLSNNHLE 223  
  
Qy 241 KNLQLLMDRVDEMDSQDIYKNTYMRNTSKQOQKHQYQORQOENMQORSGEPPLPEED 300  
Db 224 KSLQLLMDRVDELSDQIVKNTYLRNTSKQOQKHQYQORQOENMQORSGRGETSLPDED 283  
  
Qy 301 LSKLFPKPPQPPARMDSLLIAGQINTYCNQIKFTAQNLGKLFMAQALQE 349  
Db 284 VNKLFKPPVPFSPRMDSLIAGQINTYIQHIKFTSQNLGKLFMAEAMQD 332  
  
RESULT 8  
AAH64151 ID AAH64151 PRELIMINARY; PRT; 335 AA.  
AC AAH64151;  
DT 25-MAR-2004 (TrEMBLrel. 27, Created)  
DT 25-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 25-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein MGC75580.  
GN MGC75580.  
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus; Silurana.  
OX NCBI\_TaxID=8364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,







```

Q9U9Q4
ID Q9U9Q4: Q9VMU9; PRELIMINARY; PRT; 338 AA.
AC Q9U9Q4: Q9VMU9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TRENBLrel. 27, Last annotation update)
DE CG9124-PA (CG9124-pb)
GN Name=eIF-3p40; ORFNames=CG9124;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannk C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibgweg M.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laško P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Furi V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Swirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:

```

```

a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE003608; AAF52210.3; -.
DR IntAct; Q9U9Q4; -.
DR FlyBase; FBgn0022023; eIF-3p40.
DR InterPro; IPR003639; Mov34-1.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD363422; Mov34-1; 1.
DR SMART; SM00232; JAB MPN; 1.
DR SQU SEQUENCE 338 AA; 38407 MW; 76B375DCC2FDA732 CRC64;

```

```

Query Match 42.5%; Score 771; DB 2; Length 338;
Best Local Similarity 45.2%; Pred. No. 4.7e-46;
Matches 149; Conservative 70; Mismatches 109; Indels 2; Gaps 1;
QY 19 AGAAGKGGKGGGSAVKQVQIDGLVVLKTKHVEEGQGTVEVQVGLGLVVDRLRI 78
Db 2 ANRANRHAARTEDSDNTINYVOCGLAVMKVHKCHEESNMDDLAQGLGLVVDKCLSI 61
QY 79 TNCFFPPQHTDDADFDEVQYQMEWRMSLRHVNIDHLHVGWYQSTYVGSFVTRALDSQF 138
Db 62 TNCFFPPK--SGDETMDDEMYQLTWRLRRVNVVDHLHVGWYQSSDVGNLSMALLESY 119
QY 139 SYQHAIESVLLYDPIKTAQGLSLKAYRLTPKLMVECKEDFSPALKKANITFYMF 198
Db 120 HYQTSLIESVWVYDTPKSSRGFLCLKAYRLTPQAIQMYKOGDFTPEAFRTLKVGYENLF 179
QY 199 EEVPIVKNSHLINVLMMWELEKSAVADKHELLSLASSNHLKQLQLMDRVDEMSSQDIV 258
Db 180 AIPVIVKNSPLTNIMMSLNELLPEDKGNFLDLGTATVLENQMSRLIERVDELYQEA 239
QY 259 KYNTWNTSKQQQKHQYQRRQOENQORSGRGPPLPEEDLSKLFKPPQPPARMDSLL 318
Db 240 RYNYQVVFVKQDTEKRALAKLAENAVRTSKGPTVPEEVIKQFPMTPAPNRLTATI 299
QY 319 INQINTYCNKIKETQNIQGLFWAQALQ 348
Db 300 TSGQINTHAQIAQFCQSLSLAKLFIITESLQ 329

```

```

RESULT 14
AAFS2210
ID AAF52210 PRELIMINARY; PRT; 338 AA.
AC AAF52210;
DT 01-APR-2004 (TRENBLrel. 27, Created)
DT 01-APR-2004 (TRENBLrel. 27, Last sequence update)
DE 01-APR-2004 (TRENBLrel. 27, Last annotation update)
DE CG9124-PA.
GN eIF-3p40 OR CG9124.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```



OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.W., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams J.M., Woodager, Worley J.S., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL SCIENCE 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Ceiniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
 RT *melanogaster* euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Ceiniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,

RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Ceiniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003608; AAF52210.3; -;  
 DR FlyBase; FBgn0022023; eIF-3p40.  
 SQ SEQUENCE 338 AA; 38407 MW; 76E375DCC2FDA732 CRC64;  
 Query Match 42.5%; Score 771; DB 2; Length 338;  
 Best Local Similarity 45.2%; Pred. No. 4.7e-46;  
 Matches 149; Conservative 70; Mismatches 109; Indels 2; Gaps 1;  
 Qy 19 AGAAGKGGKGGSDSAVKQVQIDGLVVLTKIKHQERGQCTEVVQVGLLVVDEDLRI 78  
 Db 2 ANRANRHAARTEDSDNTINYQCDGLAVMRKMKCHBESSNNDLAQGALLGLVVDKCLEI 61  
 Qy 79 TNCFPFPFOTEDDADFOEVQYQMEMRSLRHVNIDHLHVGVYQSTYYSFVTRALLDSQF 138  
 Db 62 TNCFPFPK--SGDETMDDEMYQLVMRRLRVVDHLHVGVYQSDVCGNSLSMALESQY 119  
 Qy 139 SYQHAIBESVLLIYDPIKTAQGSLSLKAYRLTPKLMVEVCKSKDPSPEALKKANITFEVMP 198  
 Db 120 HYQTSIEESVVVVYDTOKSSRGFLCLKAYRLTPQAQIMYKDGDFTEAFRLTKVGYENLF 179  
 Qy 199 EEVPIVTKNSHLINVLWMELEKSAVADKKHELLSLASSNHLKGLQLLLMDRVDMSQDIV 258  
 Db 180 AEIPIVINKSPITNIMSELNELLPEKGNHFLDLGTATVLENQMSIERVDELYQEA 239  
 Qy 259 KYNTYMRNTSKQQQKHQYQQRQEQNMQRSGEPPLPEEDLSKLFKPPPOPPARMDSLL 318  
 Db 240 RYNYQYQVVFQDTEKHRLAKLAENAVRTSKGPTVPEEVIKQFPFMTAPNRLTATI 299  
 Qy 319 IAGQINYCNIKFETQNLKLFMAQALQ 348  
 Db 300 TSGQINTHAQIAFCQSQSLAKLFITESLQ 329  
 RESULT 15  
 ID Q6Y1Z6 PRELIMINARY; PRT; 149 AA.  
 AC Q6Y1Z6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Translation initiation factor 3 (Fragment).  
 OS *Pagrus major* (Red sea bream) (Chrysophrys major).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Sparidae; Pagrus.  
 OX NCBI\_TaxID=143350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RA Chen S.L., Xu M.Y.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY190743; AAP20218.1; -;  
 DR GO; GO:0003743; F:translation initiation factor activity; IBA.  
 KW Initiation factor.  
 FT NON TER 149  
 SQ SEQUENCE 149 AA; 17246 MW; 26DA91A07DF0CB79 CRC64;  
 Query Match 38.5%; Score 697; DB 2; Length 149;  
 Best Local Similarity 89.9%; Pred. No. 2.6e-41;

	Matches	134;	Conservative	8;	Mismatches	7;	Indels	0;	Gaps	0;
Qy	101	MEMRSRHHVNIDHHLHVGWYOSTYYGSPVTRALLDSQFSYOHAIIEESVVLIIYDPIKTAQG	160							
Db	1	MEMRSRHHVNIDHHLHVGWYOSTYYGSPVTRALLDSQFSYOHAIIEESVVLIIYDPIKTAQG	60							
Qy	161	SLSLKAYRLTPKLMVEVCKEKFSPKALKNITFEYMFEEVPIVINKSHLINVLMMWELEK	220							
Db	61	SLSLKAYRLTPKLMVEVCKEKFSPKALKNITFEYMFEEVPIVINKSHLINVLMMWELEK	120							
Qy	221	KSADKXHELLSLASSNHLGNLQLLMDR	249							
Db	121	KSTVADKHELLNLSSNNHLEKSLQLLMDR	149							

Search completed: December 30, 2004, 20:50:46  
Job time : 416.179 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:34:10 ; Search time 2.06297 Seconds  
(without alignments)  
289.321 Million cell updates/sec

Title: US-10-017-327-3  
Perfect score: 44  
Sequence: 1 FLQLLMEPV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	77.3	318	4	US-09-583-110-2976 Sequence 2976, Ap
2	33	75.0	305	4	US-09-710-279-1264 Sequence 1264, Ap
3	33	75.0	324	3	US-09-134-001C-5525 Sequence 5525, Ap
4	33	75.0	368	3	US-09-413-574-4 Sequence 4, Appli
5	33	75.0	409	4	US-09-538-092-1214 Sequence 1214, Ap
6	33	75.0	2802	3	US-09-542-331-1 Sequence 1, Appli
7	33	75.0	2802	3	US-09-510-791-1 Sequence 1, Appli
8	32	72.7	107	4	US-09-328-352-5911 Sequence 5811, Ap
9	32	72.7	195	4	US-09-134-000C-4338 Sequence 4338, Ap
10	32	72.7	323	4	US-09-107-532A-5460 Sequence 5460, Ap
11	32	72.7	406	4	US-09-248-796A-17915 Sequence 17915, A
12	32	72.7	434	4	US-09-005-286B-2 Sequence 2, Appli
13	32	72.7	642	1	US-08-706-936-3 Sequence 3, Appli
14	31	70.5	131	4	US-09-270-767-37595 Sequence 37595, A
15	31	70.5	131	4	US-09-270-767-52812 Sequence 52812, A
16	31	70.5	375	3	US-09-252-149B-35 Sequence 35, Appli
17	31	70.5	375	4	US-09-451-501-27 Sequence 27, Appli
18	31	70.5	375	4	US-09-686-344-27 Sequence 27, Appli
19	31	70.5	375	4	US-09-626-896-18 Sequence 18, Appli
20	31	70.5	375	4	US-09-485-046-14 Sequence 14, Appli
21	31	70.5	1000	3	US-09-193-562D-30 Sequence 30, Appli
22	31	70.5	1000	4	US-10-055-412B-30 Sequence 30, Appli
23	31	70.5	1253	1	US-08-252-966B-12 Sequence 12, Appli
24	31	70.5	1261	1	US-08-252-966B-18 Sequence 18, Appli
25	30	68.2	103	4	US-09-302-626B-38 Sequence 38, Appli
26	30	68.2	116	4	US-09-252-991A-29715 Sequence 29715, A
27	30	68.2	327	4	US-09-252-991A-33067 Sequence 33067, A

28	30	68.2	475	4	US-09-248-796A-18640 Sequence 18640, A
29	30	68.2	515	4	US-09-434-840-63 Sequence 63, Appli
30	30	68.2	526	4	US-09-434-840-2 Sequence 2, Appli
31	30	68.2	540	4	US-09-252-991A-18305 Sequence 18305, A
32	30	68.2	541	4	US-09-434-840-55 Sequence 55, Appli
33	30	68.2	728	3	US-08-915-337-2 Sequence 2, Appli
34	30	68.2	1332	4	US-09-252-991A-25772 Sequence 25772, A
35	30	68.2	1501	4	US-09-710-279-2850 Sequence 2850, Ap
36	30	68.2	1529	3	US-09-134-001C-3945 Sequence 3945, Ap
37	29	65.9	17	4	US-09-641-803-2 Sequence 2, Appli
38	29	65.9	44	3	US-09-046-894-38 Sequence 38, Appli
39	29	65.9	45	1	US-08-480-784-83 Sequence 83, Appli
40	29	65.9	45	1	US-08-483-553-83 Sequence 83, Appli
41	29	65.9	45	1	US-08-487-002-83 Sequence 83, Appli
42	29	65.9	45	1	US-08-483-554B-83 Sequence 83, Appli
43	29	65.9	45	1	US-08-488-011B-83 Sequence 83, Appli
44	29	65.9	45	3	US-08-850-727-83 Sequence 83, Appli
45	29	65.9	45	5	PCT-US95-10202-83 Sequence 83, Appli

ALIGNMENTS

RESULT 1

US-09-583-110-2976  
; Sequence 2976, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PTH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 2976  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-2976

Query Match 77.3%; Score 34; DB 4; Length 318;  
Best Local Similarity 75.0%; Pred. No. 37;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLLMEP 8  
Db 152 FLQVLMMP 159  
|||:|:|

RESULT 2

US-09-710-279-1264  
; Sequence 1264, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 1264  
; LENGTH: 305  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-1264

Query Match 75.0%; Score 33; DB 4; Length 305;  
Best Local Similarity 75.0%; Pred. No. 57;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLMEP 8  
|||||  
Db 138 FLQLMNP 145

## RESULT 3

US-09-134-001C-5525  
; Sequence 5525, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134.001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5525  
; LENGTH: 324  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5525

Query Match 75.0%; Score 33; DB 3; Length 324;  
Best Local Similarity 75.0%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLMEP 8  
|||||  
Db 157 FLQLMNP 164

## RESULT 4

US-09-413-574-4  
; Sequence 4, Application US/09413574  
; Patent No. 6235972  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; APPLICANT: Tagliani, Laura  
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof  
; FILE REFERENCE: 0964  
; CURRENT APPLICATION NUMBER: US/09/413,574  
; CURRENT FILING DATE: 1999-10-06  
; EARLIER APPLICATION NUMBER: 60/109,728  
; EARLIER FILING DATE: 1998-11-23  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-413-574-4

Query Match 75.0%; Score 33; DB 3; Length 368;  
Best Local Similarity 87.5%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLMEP 8  
|||||

Db 291 FLQLMEP 298

## RESULT 5

US-09-538-092-1214  
; Sequence 1214, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapatSeqFormatter Version 0.9  
; SEQ ID NO 1214  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number P54727  
US-09-538-092-1214

Query Match 75.0%; Score 33; DB 4; Length 409;  
Best Local Similarity 66.7%; Pred. No. 78;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLMEPV 9  
|:|:|  
Db 324 FIQLMEPV 332

## RESULT 6

US-09-542-331-1  
; Sequence 1, Application US/09542331  
; Patent No. 6261761  
; GENERAL INFORMATION:  
; APPLICANT: Zhong, Yi  
; APPLICANT: Guo, Hui-Fu  
; TITLE OF INVENTION: NF1 Protein and Its Role in Activation  
; TITLE OF INVENTION: of Adenylyl Cyclase by PACAP38-Like Neuropeptides  
; FILE REFERENCE: 1314.1047003  
; CURRENT APPLICATION NUMBER: US/09/542,331  
; CURRENT FILING DATE: 2000-04-04  
; EARLIER APPLICATION NUMBER: US 09/046,745  
; EARLIER FILING DATE: 1998-03-24  
; EARLIER APPLICATION NUMBER: US 60/041,469  
; EARLIER FILING DATE: 1997-03-24  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2802  
; TYPE: PRT  
; ORGANISM: Drosophila  
US-09-542-331-1

Query Match 75.0%; Score 33; DB 3; Length 2802;  
Best Local Similarity 55.6%; Pred. No. 5.8e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLMEPV 9  
:|:|:  
Db 1340 YLQMLEPL 1348

## RESULT 7

```
US-09-510-791-1
; Sequence 1, Application US/09510791
; Patent No. 6365126
; GENERAL INFORMATION:
; APPLICANT: Zhong, Yi
; APPLICANT: Guo, Hui-Fu
; APPLICANT: Tong, Jiayuan
; TITLE OF INVENTION: Improvement of Learning and Short Term
; TITLE OF INVENTION: Memory Defects with Neurofibromatosis 1(NF1) Expression
; FILE REFERENCE: 1314.1047-002
; CURRENT APPLICATION NUMBER: US/09/510,791
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: US 09/046,745
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: US 60/041,469
; PRIOR FILING DATE: 1997-03-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2802
; TYPE: PRT
; ORGANISM: Drosophila
US-09-510-791-1
Query Match 75.0%; Score 33; DB 3; Length 2802;
Best Local Similarity 55.6%; Pred. No. 5.8e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9
Db 1340 YLQMLLEPL 1348

RESULT 8
US-09-328-352-5811
; Sequence 5811, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5811
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5811
Query Match 72.7%; Score 32; DB 4; Length 107;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9
Db 2 YLEKLMEPV 10

RESULT 9
US-09-134-000C-4338
; Sequence 4338, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4338
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4338
Query Match 72.7%; Score 32; DB 4; Length 195;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 8
Db 27 FLQVIMDP 34

RESULT 10
US-09-107-532A-5460
; Sequence 5460, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5460:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 323 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...323
; SEQUENCE DESCRIPTION: SEQ ID NO: 5460:
US-09-107-532A-5460
Query Match 72.7%; Score 32; DB 4; Length 323;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 8
```

```
Db      155 FLQVIMDP 162
|||:::|
; APPLICANT: Keich Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17915
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (403)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-17915

Query Match      72.7%; Score 32; DB 4; Length 406;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 FLQLMEPV 9
|||:::|
Db      38 FLQLLSHPI 46

US-09-005-286B-2
; Sequence 2, Application US/09005286B
; Patent No. 6756491
; GENERAL INFORMATION:
; APPLICANT: EVANS, RONALD M.
; TITLE OF INVENTION: NOVEL STEROID-ACTIVATED NUCLEAR RECEPTORS AND USES
; FILE REFERENCE: 088802-5201
; CURRENT APPLICATION NUMBER: US/09/005,286B
; CURRENT FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-005-286B-2

Query Match      72.7%; Score 32; DB 4; Length 434;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FLQLMEPV 9
|||:::|
Db      315 FQQLLEPM 323

US-08-706-936-3
; Sequence 3, Application US/08706936
; Patent No. 5792851
; GENERAL INFORMATION:
```

```
; APPLICANT: VICTOR L. SCHUSTER AND RUN LU
; TITLE OF INVENTION: HUMAN PROSTAGLANDIN TRANSPORTER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,936
; FILING DATE: SEPTEMBER 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG J. ARNOLD
; REGISTRATION NUMBER: 34,287
; REFERENCE/DOCKET NUMBER: 96700/406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: PROTEIN
; DESCRIPTION: YES
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: RAT
; INDIVIDUAL ISOLATE: PROSTAGLANDIN TRANSPORTER
US-08-706-936-3

Query Match      72.7%; Score 32; DB 1; Length 642;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FLQLMEPV 9
|||:::|
Db      317 FLRLMNPL 325

RESULT 14
US-09-270-767-37595
; Sequence 37595, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 37595
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-37595

Query Match      70.5%; Score 31; DB 4; Length 131;
Best Local Similarity 44.4%; Pred. No. 60;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FLQLMEPV 9
```

```
Db          90 FVQILVQPI 98
|:|:|:|:|
RESULT 15
US-09-270-767-52812
; Sequence 52812, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52812
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-52812
Query Match      70.5%; Score 31; DB 4; Length 131;
Best Local Similarity 44.4%; Pred. NO. 60;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy      1 FLQLMEPV 9
|:|:|:|:|
Db          90 FVQILVQPI 98

Search completed: December 30, 2004, 20:53:54
Job time : 2.06297 secs
```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:35:11 ; Search time 8.81864 Seconds  
(without alignments)  
367.126 Million cell updates/sec

Title: US-10-017-327-3

Perfect score: 44

Sequence: 1 FLQLMEPV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	11	US-09-870-216C-3
2	44	100.0	9	13	US-10-017-327-3
3	35	79.5	67	9	US-09-764-869-1075
4	35	79.5	67	14	US-10-091-504-1075
5	35	79.5	67	15	US-10-227-577-1075
6	35	79.5	329	16	US-10-478-551-2
7	35	79.5	363	15	US-10-424-599-219691
8	35	79.5	363	15	US-10-424-599-219691
9	34	77.3	79	17	US-10-425-115-349203
10	34	77.3	180	16	US-10-437-963-114660
11	34	77.3	274	16	US-10-437-963-132856
12	34	77.3	307	9	US-09-799-848-3
13	34	77.3	307	11	US-09-781-712B-3

14	34	77.3	307	17	US-10-679-761-8	Sequence 8, Appli
15	34	77.3	318	16	US-10-415-478A-5	Sequence 5, Appli
16	34	77.3	504	15	US-10-425-114-45949	Sequence 45949, A
17	34	77.3	770	16	US-10-437-963-163658	Sequence 163658,
18	34	77.3	798	17	US-10-425-115-200455	Sequence 200455,
19	33	75.0	80	17	US-10-425-115-318997	Sequence 318997,
20	33	75.0	117	11	US-09-864-408A-7182	Sequence 7182, Ap
21	33	75.0	138	17	US-10-425-115-223947	Sequence 223947,
22	33	75.0	174	17	US-10-425-115-223948	Sequence 223948,
23	33	75.0	193	16	US-10-767-701-35283	Sequence 35283, A
24	33	75.0	349	17	US-10-425-115-223949	Sequence 223949,
25	33	75.0	368	9	US-09-805-550-4	Sequence 4, Appli
26	33	75.0	1224	16	US-10-437-963-138501	Sequence 138501,
27	32	72.7	73	16	US-10-425-115-307557	Sequence 307557,
28	32	72.7	73	16	US-10-767-701-56147	Sequence 56147, A
29	32	72.7	79	17	US-10-425-115-358291	Sequence 358291,
30	32	72.7	104	16	US-10-767-701-62832	Sequence 62832, A
31	32	72.7	169	17	US-10-425-115-360733	Sequence 360733,
32	32	72.7	189	15	US-10-424-599-245981	Sequence 245981,
33	32	72.7	198	15	US-10-424-599-196742	Sequence 196742,
34	32	72.7	269	15	US-10-424-599-234629	Sequence 234629,
35	32	72.7	272	15	US-10-425-114-53064	Sequence 53064, A
36	32	72.7	276	17	US-10-425-115-367551	Sequence 367551,
37	32	72.7	330	17	US-10-775-678-54	Sequence 54, Appli
38	32	72.7	334	16	US-10-437-963-108157	Sequence 108157,
39	32	72.7	362	17	US-10-739-930-9255	Sequence 9255, Ap
40	32	72.7	420	17	US-10-425-115-365824	Sequence 365824,
41	32	72.7	434	10	US-09-143-828-2	Sequence 2, Appli
42	32	72.7	434	10	US-09-227-718-2	Sequence 2, Appli
43	32	72.7	434	10	US-09-840-008-2	Sequence 2, Appli
44	32	72.7	434	14	US-10-081-555C-2	Sequence 2, Appli
45	32	72.7	434	15	US-10-226-997-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-870-216C-3  
; Sequence 3, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 69126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-216C-3

Query Match 100.0%; Score 44; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLQLMEPV 9  
| | | | | | | |  
Db 1 FLQLMEPV 9

RESULT 2  
US-10-017-327-3  
; Sequence 3, Application US/10017327

```
; Publication No. US20020155471A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
; FILE OF INVENTION: METHODS FOR USING SAME
; FILE REFERENCE: G2 2101.20
; CURRENT APPLICATION NUMBER: US/10/017,327
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-327-3

Query Match      100.0%; Score 44; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9
Db 1 FLQLLMEPV 9

RESULT 3
US-09-764-869-1075
; Sequence 1075, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1075
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-1075

Query Match      79.5%; Score 35; DB 9; Length 67;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9
Db 7 FLQLLTPV 15

RESULT 4
US-10-091-504-1075
; Sequence 1075, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1075
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-227-577-1075

Query Match      79.5%; Score 35; DB 14; Length 67;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9
Db 7 FLQLLTPV 15

RESULT 5
US-10-227-577-1075
; Sequence 1075, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1075
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-227-577-1075

Query Match      79.5%; Score 35; DB 15; Length 67;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 FLOLLMEPV 9  
||||: ||  
Db 7 FLOLLTPV 15

RESULT 6

US-10-478-551-2  
; Sequence 2, Application US/10478551  
; Publication No. US2004017181A1  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham Corporation  
; APPLICANT: Steven Anthony Klierer  
; APPLICANT: Jodi Marie Maglich  
; APPLICANT: John Tomlin Moore  
; APPLICANT: Linda Becker Moore  
; APPLICANT: Timothy Mark Willson  
; TITLE OF INVENTION: NONHUMAN PREGNANE X RECEPTOR SEQUENCES  
; TITLE OF INVENTION: FOR USE IN COMPARITIVE PHARMACOLOGY  
; FILE REFERENCE: P03855W0  
; CURRENT APPLICATION NUMBER: US/10/478,551  
; CURRENT FILING DATE: 2003-11-21  
; PRIOR APPLICATION NUMBER: 60/293,380  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 329  
; TYPE: PRT  
; ORGANISM: Canine  
US-10-478-551-2

Query Match 79.5%; Score 35; DB 16; Length 329;  
Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
||||: |||  
Db 210 FQOLLLEPV 218

RESULT 7

US-10-424-599-170467  
; Sequence 170467, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 170467  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(363)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_124948C.1.pep  
US-10-424-599-170467

Query Match 79.5%; Score 35; DB 15; Length 363;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
||||: |||

Db 286 FLOLINEPV 294

RESULT 8

US-10-424-599-219691  
; Sequence 219691, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 219691  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_40409C.1.pep  
US-10-424-599-219691

Query Match 79.5%; Score 35; DB 15; Length 363;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
||||: |||  
Db 286 FLOLINEPV 294

RESULT 9

US-10-425-115-349203  
; Sequence 349203, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 349203  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_81639C.1.pep  
US-10-425-115-349203

Query Match 77.3%; Score 34; DB 17; Length 79;  
Best Local Similarity 66.7%; Pred. No. 42;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
:||||: |||  
Db 66 YLQLLRPV 74

RESULT 10

US-10-437-963-114660  
; Sequence 114660, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 114660  
LENGTH: 180  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(180)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_18329C.1.pap  
US-10-437-963-114660

Query Match 77.3%; Score 34; DB 16; Length 180;  
Best Local Similarity 77.8%; Pred. No. 99;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9  
| | | | | | |  
Db 169 FCQLLMHPV 177

## RESULT 11

US-10-437-963-132856  
Sequence 132856, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 132856  
LENGTH: 274  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(274)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_34784C.1.pap  
US-10-437-963-132856

Query Match 77.3%; Score 34; DB 16; Length 274;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9  
| | | | | | |  
Db 184 FCQLLMHPV 192

RESULT 12  
US-09-799-848-3  
Sequence 3, Application US/09799848  
Patent No. US20010044145A1  
GENERAL INFORMATION:  
APPLICANT: Monia, Brett  
APPLICANT: Cook, Phillip  
APPLICANT: Crooke, Stanley  
APPLICANT: Wu, Hongjiang  
APPLICANT: Lima, Walter  
TITLE OF INVENTION: METHODS OF USING MAMMALIAN RNASE H AND COMPOSITIONS THEREOF  
FILE REFERENCE: ISPH-0521  
CURRENT APPLICATION NUMBER: US/09/799,848  
CURRENT FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: US 09/343,809  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 09/684,254  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: US 09/203,716  
PRIOR FILING DATE: 1998-12-02  
PRIOR APPLICATION NUMBER: US 60/067,458  
PRIOR FILING DATE: 1997-12-04  
PRIOR APPLICATION NUMBER: US 09/453,514  
PRIOR FILING DATE: 1999-12-01  
PRIOR APPLICATION NUMBER: US 09/144,611  
PRIOR FILING DATE: 1998-08-31  
PRIOR APPLICATION NUMBER: US 08/861,306  
PRIOR FILING DATE: 1997-04-21  
PRIOR APPLICATION NUMBER: US 08/244,993  
PRIOR FILING DATE: 1994-06-21  
PRIOR APPLICATION NUMBER: US 07/814,961  
PRIOR FILING DATE: 1991-12-24  
PRIOR APPLICATION NUMBER: US 09/462,280  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: PCT/US98/13966  
PRIOR FILING DATE: 1998-07-06  
PRIOR APPLICATION NUMBER: US 08/889,296  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: US 08/411,734  
PRIOR FILING DATE: 1995-04-03  
PRIOR APPLICATION NUMBER: US 08/007,996  
PRIOR FILING DATE: 1993-10-21  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 307  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-799-848-3

Query Match 77.3%; Score 34; DB 9; Length 307;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9  
| | | | | | |  
Db 227 FLQLSVEPV 235

## RESULT 13

US-09-781-712B-3  
Sequence 3, Application US/09781712B  
Publication No. US20040180433A1  
GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T  
APPLICANT: Lima, Walter  
APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Methods of Using Mammalian RNase H and Compositions Thereof  
FILE REFERENCE: ISPH-0520  
CURRENT APPLICATION NUMBER: US/09/781,712B  
CURRENT FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: US 60/067,458

Query Match 77.3%; Score 34; DB 16; Length 274;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLQLLMEPV 9  
| | | | | | |  
Db 184 FCQLLMHPV 192

; PRIOR FILING DATE: 1997-12-04  
; PRIOR APPLICATION NUMBER: US 09/203,716  
; PRIOR FILING DATE: 1998-12-02  
; PRIOR APPLICATION NUMBER: US 09/343,809  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: US 09/684,254  
; PRIOR FILING DATE: 2000-10-06  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-781-712B-3

Query Match 77.3%; Score 34; DB 11; Length 307;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 9  
||||:||||  
Db 227 FLOLSVEPV 235

RESULT 14  
US-10-679-761-8  
; Sequence 8, Application US/10679761  
; Publication No. US20040248145A1  
; GENERAL INFORMATION:  
; APPLICANT: Isis Pharmaceuticals, Inc.  
; APPLICANT: Crooke, Stanley T.  
; APPLICANT: Lima, Walter  
; APPLICANT: Wu, Hongjiang  
; TITLE OF INVENTION: Methods of Using Mammalian RNase H and Compositions Thereof  
; FILE REFERENCE: ISPH-0790  
; CURRENT APPLICATION NUMBER: US/10/679,761  
; CURRENT FILING DATE: 2003-10-06  
; PRIOR APPLICATION NUMBER: US 10/358,439  
; PRIOR FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: US 09/992,738  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/781,712  
; PRIOR FILING DATE: 2001-02-12  
; PRIOR APPLICATION NUMBER: US 09/861,205  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 09/684,254  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 09/343,809  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: US 09/203,716  
; PRIOR FILING DATE: 1998-12-02  
; PRIOR APPLICATION NUMBER: US 60/067,458  
; PRIOR FILING DATE: 1997-12-04  
; PRIOR APPLICATION NUMBER: US 60/248,950  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: US 60/497,412  
; PRIOR FILING DATE: 2003-08-21  
; NUMBER OF SEQ ID NOS: 104  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 8  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-679-761-8

Query Match 77.3%; Score 34; DB 17; Length 307;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 9  
||||:||||  
Db 227 FLOLSVEPV 235

RESULT 15  
US-10-415-478A-5  
; Sequence 5, Application US/10415478A  
; Publication No. US20040116661A1  
; GENERAL INFORMATION:  
; APPLICANT: Jeremy Stuart Brown  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Streptococcal Genes  
; FILE REFERENCE: GJE-6571  
; CURRENT APPLICATION NUMBER: US/10/415.478A  
; CURRENT FILING DATE: 2003-12-29  
; PRIOR APPLICATION NUMBER: PCT/GB01/04749  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: 0026231.1  
; PRIOR FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 00028345.7  
; PRIOR FILING DATE: 2000-11-21  
; PRIOR APPLICATION NUMBER: 0102666.5  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/288,118  
; PRIOR FILING DATE: 2001-05-02  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Salmonella pneumoniae  
US-10-415-478A-5

Query Match 77.3%; Score 34; DB 16; Length 318;  
Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLMPE 8  
||||:||||  
Db 152 FLOVIMDP 159

Search completed: December 30, 2004, 21:00:31  
Job time : 10.8186 secs

**This Page Blank (uspto)**





CC The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen eIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents synthetic epitope 1 of human cancer antigen eIF3

XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
 |||||  
 Db 1 FLQLMEPV 9

# RESULT 2

ID ABR82213 standard; peptide; 9 AA.

AC ABR82213;

DT 13-OCT-2003 (first entry)

DE Human antigen eIF3 derived compound 1.

KW Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
 KW cytostatic; gene therapy; human; antigen.

XX Synthetic.

OS Homo sapiens.

XX WO2003050543-A1.

XX 19-JUN-2003.

PF 05-DEC-2001; 2001WO-US047997.

XX 05-DEC-2001; 2001WO-US047997.

XX (GENZ ) GENZYME CORP.

PI Nicolette CA;

XX WPI; 2003-532936/50.

DR N-PSDB; ACC85030.

XX Aiding in the diagnosis of a neoplastic condition, useful for treating  
 PT cancer and related malignancies comprises determining the amount of  
 FT expression of an eIF3 protein in a test sample isolated from the cell or  
 PT tissue.

XX Claim 12; Page 30; 77pp; English.

PS The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic  
 CC condition or susceptibility to a neoplastic condition based on the amount

CC of expression of the eIF3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. Sequences ABR82213-16  
 CC represent compounds derived from the human antigen eIF3

SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9  
 |||||  
 Db 1 FLQLMEPV 9

# RESULT 3

ABR808366

ID ABR808366 standard; protein; 352 AA.

XX ABR808366;

DT 07-MAY-2002 (first entry)

DE Human cancer antigen eIF3 variant 1 amino acid sequence.

XX Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 242 /note= "wild-type Asn is replaced by Phe"

FT Domain 242 /note= "HLA-2 binding residue"

FT Domain 243 /note= "HLA-2 binding residue"

FT Domain 244.249 /note= "T-cell receptor (TCR) binding domain"

FT Misc-difference 248 /note= "wild-type Asp is replaced by Glu"

FT Misc-difference 249 /note= "wild-type Arg is replaced by Pro"

FT Domain 250 /note= "HLA-2 binding residue"

XX WO200192307-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017456.

XX 31-MAY-2000; 2000US-0209391P.

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

XX (GENZ ) GENZYME CORP.

PI Nicolette CA;

XX WPI; 2002-139606/18.

XX New therapeutic compounds useful against human ovarian cancer, for  
 PT modulating immune response in a subject, and for generating antibodies  
 FT that specifically recognize and bind to these molecules.

PS Claim 6; Page; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory

CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen E1F3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human cancer antigen e1f3 variant 1 amino acid sequence.  
 CC Note: This sequence is not present in the specification, but may be  
 CC created from the sequence of the wild-type human cancer antigen e1f3  
 CC sequence given in ABB08360

XX Sequence 352 AA;

Query Match 100.0%; Score 44; DB 5; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.1; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 9

Db 242 FLQLMEPV 250

RESULT 4

AAU22301

ID AAU22301 standard; protein; 67 AA.

XX AC

XX AAU22301;

XX DT 18-DEC-2001 (first entry)

XX DE

XX Human cardiovascular system antigen polypeptide SEQ ID No 1075.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
 KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
 KW antineumatic; antiproliferative; cytostatic; cardiac; neuroprotective;  
 KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
 KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
 KW hyperproliferative disorder; breast; liver; cardiovascular disorder;  
 KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
 KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
 KW gastrointestinal disorder; renal disorder; respiratory disorder;  
 KW wound healing; skin aging; organ transplantation; tissue regeneration;  
 KW anti-infertility.

XX OS Homo sapiens.

XX WO200155321-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001340.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 19-MAY-2000; 2000US-0198123P.

XX PR 07-JUN-2000; 2000US-0205515P.

XX PR 28-JUN-2000; 2000US-0209467P.

XX PR 30-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.



PR	14-AUG-2000;	2000US-02252666P;	PR
PR	14-AUG-2000;	2000US-0225267P;	PR
PR	14-AUG-2000;	2000US-0225268P;	PR
PR	14-AUG-2000;	2000US-0225270P;	PR
PR	14-AUG-2000;	2000US-0225270P;	PR
PR	14-AUG-2000;	2000US-0225447P;	PR
PR	14-AUG-2000;	2000US-0225575P;	PR
PR	14-AUG-2000;	2000US-0225578P;	PR
PR	14-AUG-2000;	2000US-0225579P;	PR
PR	18-AUG-2000;	2000US-0226277P;	PR
PR	22-AUG-2000;	2000US-0226681P;	PR
PR	22-AUG-2000;	2000US-0227182P;	PR
PR	22-AUG-2000;	2000US-0227182P;	PR
PR	23-AUG-2000;	2000US-0227009P;	PR
PR	30-SEP-2000;	2000US-0228924P;	PR
PR	01-SEP-2000;	2000US-0228287P;	PR
PR	01-SEP-2000;	2000US-0229343P;	PR
PR	01-SEP-2000;	2000US-0229344P;	PR
PR	01-SEP-2000;	2000US-0229345P;	PR
PR	05-SEP-2000;	2000US-0229509P;	PR
PR	05-SEP-2000;	2000US-0229513P;	PR
PR	06-SEP-2000;	2000US-0230437P;	PR
PR	06-SEP-2000;	2000US-0230438P;	PR
PR	08-SEP-2000;	2000US-0231242P;	PR
PR	08-SEP-2000;	2000US-0231243P;	PR
PR	08-SEP-2000;	2000US-0231244P;	PR
PR	08-SEP-2000;	2000US-0231413P;	PR
PR	08-SEP-2000;	2000US-0231414P;	PR
PR	08-SEP-2000;	2000US-0232080P;	PR
PR	08-SEP-2000;	2000US-0232081P;	PR
PR	12-SEP-2000;	2000US-0231968P;	PR
PR	14-SEP-2000;	2000US-0232397P;	PR
PR	14-SEP-2000;	2000US-0232398P;	PR
PR	14-SEP-2000;	2000US-0232399P;	PR
PR	14-SEP-2000;	2000US-0232400P;	PR
PR	14-SEP-2000;	2000US-0232401P;	PR
PR	14-SEP-2000;	2000US-0232403P;	PR
PR	14-SEP-2000;	2000US-0232064P;	PR
PR	27-SEP-2000;	2000US-0235834P;	PR
PR	27-SEP-2000;	2000US-0234223P;	PR
PR	29-SEP-2000;	2000US-0235836P;	PR
PR	29-SEP-2000;	2000US-0236327P;	PR
PR	29-SEP-2000;	2000US-0234597P;	PR
PR	29-SEP-2000;	2000US-0234998P;	PR
PR	29-SEP-2000;	2000US-0236368P;	PR
PR	29-SEP-2000;	2000US-0236369P;	PR
PR	29-SEP-2000;	2000US-0236370P;	PR
PR	02-OCT-2000;	2000US-0236802P;	PR
PR	02-OCT-2000;	2000US-0237037P;	PR
PR	02-OCT-2000;	2000US-0237038P;	PR
PR	02-OCT-2000;	2000US-0237039P;	PR
PR	02-OCT-2000;	2000US-0237040P;	PR
PR	13-OCT-2000;	2000US-0239935P;	PR
PR	13-OCT-2000;	2000US-0239937P;	PR
PR	20-OCT-2000;	2000US-0240960P;	PR
PR	20-OCT-2000;	2000US-0241221P;	PR
PR	20-OCT-2000;	2000US-0241785P;	PR
PR	20-OCT-2000;	2000US-0241786P;	PR
PR	20-OCT-2000;	2000US-0241787P;	PR
PR	20-OCT-2000;	2000US-0241808P;	PR
PR	20-OCT-2000;	2000US-0241809P;	PR
PR	20-OCT-2000;	2000US-0241826P;	PR
PR	01-NOV-2000;	2000US-0244617P;	PR
PR	08-NOV-2000;	2000US-0246474P;	PR
PR	08-NOV-2000;	2000US-0246475P;	PR
PR	08-NOV-2000;	2000US-0246476P;	PR
PR	08-NOV-2000;	2000US-0246477P;	PR
PR	08-NOV-2000;	2000US-0246478P;	PR
PR	08-NOV-2000;	2000US-0246523P;	PR
PR	08-NOV-2000;	2000US-0246524P;	PR

PR	08-NOV-2000;	2000US-0246525P;
PR	08-NOV-2000;	2000US-0246526P;
PR	08-NOV-2000;	2000US-0246527P;
PR	08-NOV-2000;	2000US-0246528P;
PR	08-NOV-2000;	2000US-0246532P;
PR	08-NOV-2000;	2000US-0246609P;
PR	08-NOV-2000;	2000US-0246610P;
PR	08-NOV-2000;	2000US-0246611P;
PR	08-NOV-2000;	2000US-0246613P;
PR	17-NOV-2000;	2000US-0249207P;
PR	17-NOV-2000;	2000US-0249208P;
PR	17-NOV-2000;	2000US-0249209P;
PR	17-NOV-2000;	2000US-0249210P;
PR	17-NOV-2000;	2000US-0249211P;
PR	17-NOV-2000;	2000US-0249212P;
PR	17-NOV-2000;	2000US-0249213P;
PR	17-NOV-2000;	2000US-0249214P;
PR	17-NOV-2000;	2000US-0249215P;
PR	17-NOV-2000;	2000US-0249216P;
PR	17-NOV-2000;	2000US-0249217P;
PR	17-NOV-2000;	2000US-0249218P;
PR	17-NOV-2000;	2000US-0249219P;
PR	17-NOV-2000;	2000US-0249244P;
PR	17-NOV-2000;	2000US-0249245P;
PR	17-NOV-2000;	2000US-0249246P;
PR	17-NOV-2000;	2000US-0249265P;
PR	17-NOV-2000;	2000US-0249297P;
PR	17-NOV-2000;	2000US-0249299P;
PR	17-NOV-2000;	2000US-0249300P;
PR	01-DEC-2000;	2000US-0250160P;
PR	01-DEC-2000;	2000US-0250319P;
PR	05-DEC-2000;	2000US-0251030P;
PR	05-DEC-2000;	2000US-0251198P;
PR	05-DEC-2000;	2000US-0256719P;
PR	06-DEC-2000;	2000US-0251479P;
PR	08-DEC-2000;	2000US-0251856P;
PR	08-DEC-2000;	2000US-0251868P;
PR	08-DEC-2000;	2000US-0251869P;
PR	08-DEC-2000;	2000US-0251989P;
PR	08-DEC-2000;	2000US-0251990P;
PR	11-DEC-2000;	2000US-0254977P;
PR	05-JAN-2001;	2001US-0259678P;
PR	17-JAN-2001;	2001US-0076486P;

CC decrease storage capabilities, fat content or other nutritional  
 CC components. This sequence represents a human cardiovascular system  
 CC related polypeptide of the invention.  
 XX  
 SQ Sequence 67 AA;  
 Query Match 79.5%; Score 35; DB 7; Length 67;  
 Best Local Similarity 77.8%; Pred. No. 15;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLQLMEPV 9  
 Db 7 FLQLLLTPV 15  
 RESULT 6  
 ABB99785  
 ID ABB99785 standard; protein; 329 AA.  
 AC ABB99785;  
 XX 24-MAR-2003 (first entry)  
 DT  
 XX Amino acid sequence of canine PXR ligand binding domain.  
 DE  
 XX Pregnane X nuclear receptor; PXR; P450 3A4 monooxygenase;  
 KW drug metabolism.  
 KW  
 XX Canis sp.  
 OS  
 XX WO200294865-A1.  
 PN  
 XX 28-NOV-2002.  
 PD  
 XX 24-MAY-2002; 2002WO-US016445.  
 PF  
 XX 24-MAY-2001; 2001US-0293380P.  
 PR  
 XX (SMIK ) SMITHLINE BEECHAM CORP.  
 XX  
 FA  
 XX Kliever SA, Maglich JM, Moore JT, Moore LB, Willson TM;  
 PI WPI; 2003-148458/14.  
 XX N-PSDB; ABZ23279.  
 DR  
 XX New pregnane X nuclear receptor polypeptide that modulate P450 3A4 levels  
 PT or activity, useful in comparative pharmacology and for selecting  
 PT appropriate pre-clinical animal models predictive of human PXR activity.  
 XX  
 XX Claim 1; Page 47-48; 56pp; English.  
 PS  
 XX The present sequence is a pregnane X nuclear receptor (PXR) ligand  
 CC binding domain. PXR polypeptides modulate P450 3A4 monooxygenase levels  
 CC or activity. P450 3A4 catalyses the metabolism of more than 60% of all  
 CC drugs that are in use, including steroids, immunosuppressive agents,  
 CC imidazole antimycotics, and macrolide antibiotics. The polypeptide is  
 CC useful in comparative pharmacology and for selecting appropriate pre-  
 CC clinical animal models predictive of human PXR activity  
 XX  
 SQ Sequence 329 AA;  
 Query Match 79.5%; Score 35; DB 6; Length 329;  
 Best Local Similarity 77.8%; Pred. No. 78;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLQLMEPV 9  
 Db 210 FQQLLETPV 218  
 RESULT 7  
 ABB53033  
 ID ABB53033 standard; protein; 307 AA.

XX ABP53033;  
 AC  
 XX 06-NOV-2002 (first entry)  
 DT  
 XX C. elegans RNase H homologous protein sequence SEQ ID NO:3.  
 DE  
 XX RNase H; antisense technology; inhibition.  
 XX  
 KW Caenorhabditis elegans.  
 OS  
 XX WO200264841-A1.  
 PN  
 XX 22-AUG-2002.  
 PD  
 XX 12-FEB-2002; 2002WO-US004243.  
 PF  
 XX 12-FEB-2001; 2001US-00781712.  
 PR  
 XX (ISIS-) ISIS PHARM INC.  
 PA  
 XX Crooke ST, Lima WF, Wu H;  
 PI WPI; 2002-657606/70.  
 DR  
 XX Use of a mammalian, particularly human, RNase H, for treating an animal  
 PT with a disease or condition associated with a human RNase H, for  
 PT inhibiting the expression of a protein, or for reducing cellular RNA via  
 PT antisense technology.  
 XX  
 XX Example 1; Fig 1; 70pp; English.  
 PS  
 XX The present invention describes a method for promoting the inhibition of  
 CC the expression of a protein comprising employing a mammalian RNase H  
 CC polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA  
 CC complex duplex occurs. Also described is a compound 8 to 50 nucleobases  
 CC in length targeted to the nucleic acid encoding the human RNase HII  
 CC polypeptide, where the compound specifically hybridises with and inhibits  
 CC the expression of a human RNase HII polypeptide. The compound, which is  
 CC an antisense oligonucleotide, is useful for inhibiting the expression of  
 CC a human RNase HII polypeptide in cells or tissues, as well as for  
 CC treating an animal with a disease or condition associated with a human  
 CC RNase HII polypeptide. The method is useful for inhibiting the expression  
 CC of a protein, particularly for reducing cellular RNA via antisense  
 CC technology. The present sequence represents a protein sequence given in  
 CC comparison with the human RNase HII protein sequence, given in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 307 AA;  
 Query Match 77.3%; Score 34; DB 5; Length 307;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLQLMEPV 9  
 Db 227 FLQLSVEPV 235  
 RESULT 8  
 AAY85935  
 ID AAY85935 standard; protein; 318 AA.  
 XX  
 AC AAY85935;  
 XX  
 XX 10-APR-2000 (first entry)  
 DT  
 XX S. pneumoniae derived protein #144.  
 DE  
 XX Treatment; prevention; disease; diagnosis; gene therapy; screening;  
 KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.  
 KW  
 XX Streptococcus pneumoniae.  
 OS

XX WO9806734-A1.  
 XX 19-FEB-1998.  
 XX  
 XX 15-AUG-1997; 97WO-US014436.  
 XX 16-AUG-1996; 96US-0024022P.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;  
 XX Stodola RK;  
 XX WPI; 1998-159452/14.  
 XX N-PSDB; AAZ96294.  
 XX Streptococcus pneumoniae proteins and related DNA - useful for screening  
 XX compounds for antibacterial activity.  
 XX Claim 5; Page 430-431; 640pp; English.  
 XX This invention describes novel isolated Streptococcus pneumoniae  
 XX polynucleotides (see AAZ96173-296494) and their encoded proteins (see  
 XX AAZ96173-296494). The DNA, vectors and host cells described in the method  
 XX of the invention are useful for the recombinant expression of the  
 XX polypeptides. The polypeptides are useful for treatment or prevention of  
 XX disease, or diagnosis of disease related to expression or activity of  
 XX such a polypeptide. They can also be used to screen for compounds which  
 XX interact with and inhibit or activate such a polypeptide. The  
 XX polypeptides (or DNA encoding them, via gene therapy) are also useful for  
 XX inducing an immunological response in a mammal. The antagonists are  
 XX useful to inhibit such bacterial polypeptides. The polypeptides are  
 XX particularly useful to identify antimicrobial compounds and antibiotics.  
 XX They are also useful to determine their role in pathogenesis of  
 XX infection, dysfunction and disease  
 XX  
 XX Sequence 318 AA;  
 Query Match 77.3%; Score 34; DB 2; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FLQLMEP 8  
 Db 152 FLQVLMDF 159  
 |||:|:|  
 RESULT 9  
 ABG61495  
 ID ABG61495 standard; protein; 318 AA.  
 XX  
 XX ABG61495;  
 AC  
 XX 07-AUG-2003 (revised)  
 DT 27-AUG-2002 (first entry)  
 XX  
 XX Iron uptake ABC transporter polypeptide #2.  
 DE  
 XX Iron uptake ABC transporter; Streptococcal iron transporter; Sit;  
 KW antimicrobial; antibacterial; veterinary treatment; passive immunisation.  
 XX  
 XX Salmonella sp.  
 OS  
 XX WO200234773-A2.  
 FN  
 XX 02-MAY-2002.  
 PD  
 XX 26-OCT-2001; 2001WO-GB004749.  
 PF  
 XX 26-OCT-2000; 2000GB-00026231.  
 XX 21-NOV-2000; 2000GB-00028345.  
 PR 02-FEB-2001; 2001GB-00002656.  
 PR

PR 02-MAY-2001; 2001US-0288118P.  
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.  
 XX  
 XX Holden DW, Brown JS;  
 FI  
 XX WPI; 2002-463352/49.  
 DR N-PSDB; ABK83158.  
 XX  
 XX Novel Streptococcus pneumoniae iron uptake ABC transporter peptide,  
 PT useful in screening assay for identifying antimicrobial drug and in  
 PT diagnostic assay for detecting streptococcal microorganism.  
 XX  
 XX Disclosure; Page 51-52; 159pp; English.  
 XX  
 XX The invention relates to a Streptococcus pneumoniae iron uptake ABC  
 CC transporter peptide called Streptococcal iron transporter (Sit) and the  
 CC polynucleotide encoding it. The sequences are useful for therapeutic or  
 CC diagnostic purposes, in screening assays for the identification of an  
 CC antimicrobial drug, and in diagnostic assays for the detection of a  
 CC streptococcal microorganism. The sequences are used for the manufacture  
 CC of a medicament for the treatment or prevention of a condition associated  
 CC with infection by S. pneumoniae or other gram positive bacteria,  
 CC preferably for veterinary treatment, and in the production of monoclonal  
 CC and polyclonal antibodies for use in passive immunisation. This sequence  
 CC represents an S. pneumoniae iron uptake ABC transporter. (Updated on 07-  
 CC AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 318 AA;  
 Qy 1 FLQLMEP 8  
 Db 152 FLQVLMDF 159  
 |||:|:|  
 Query Match 77.3%; Score 34; DB 5; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FLQLMEP 8  
 Db 152 FLQVLMDF 159  
 |||:|:|  
 RESULT 10  
 ABU02362  
 ID ABU02362 standard; protein; 318 AA.  
 XX  
 XX ABU02362;  
 AC  
 XX 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 XX  
 XX S. pneumoniae type 4 strain protein from coding region #1940.  
 DE  
 XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 KW gene therapy; vaccine.  
 XX  
 XX Streptococcus pneumoniae; type 4 strain.  
 OS  
 XX WO200277021-A2.  
 FN  
 XX 03-OCT-2002.  
 PD  
 XX 27-MAR-2002; 2002WO-IB002163.  
 PF  
 XX 27-MAR-2001; 2001GB-00007658.  
 XX  
 XX (CHIR-) CHIRON SPA.  
 FA (GENO-) INST GENOMIC RES.  
 XX  
 XX Masignani V, Tettelin H, Fraser C;  
 FI  
 XX WPI; 2003-040579/03.  
 DR N-PSDB; ABX07652.  
 XX  
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
 PT



PT useful as medicaments for treating or preventing a disease or infection  
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or  
 PT ear infection.

PS Claim 1; SEQ ID NO 3880; 56pp; English.

XX The invention relates to a protein comprising or having at least 50%  
 CC identity to any of the 2469 amino acid sequences, identified in the  
 CC specification (available on a computer readable format), or its fragment,  
 CC expressed from 2469 identified DNA coding regions from the  
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as  
 CC AB956454. Also included are an antibody which binds one of the proteins,  
 CC treating a patient by administering the protein, DNA or antibody (in a  
 CC composition), a kit comprising first and second primers, which are the  
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a  
 CC sequence not defined in the specification, for amplifying a target  
 CC sequence contained within a Streptococcus nucleic acid sequence, where  
 CC the first primer is substantially complementary to the target sequence  
 CC and the second primer is substantially complementary to the complement of  
 CC the target sequence, and where the parts of the primers having  
 CC substantial complementarity define the termini of the target sequence to  
 CC be amplified, assay comprising contacting a test compound with the  
 CC protein, and determining whether the test compound binds to the protein  
 CC and a Streptococcus pneumoniae bacterium, where one or more genes  
 CC encoding the proteins has been rendered inactive. The proteins, nucleic  
 CC acid molecules, antibody and compositions are useful as medicaments for  
 CC treating or preventing a disease or infection due to streptococcus  
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis  
 CC media or ear infection. They are also useful in developing vaccines,  
 CC diagnostics and antibiotics. The methods are useful for identifying  
 CC immunodominant proteins. The present sequence is one of the 2469 proteins  
 CC expressed by the identified coding regions from the genomic sequence.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-OCT-2003 to  
 CC standardise OS field)

XX Sequence 318 AA;

Query Match 77.3%; Score 34; DB 6; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLMPEP 8  
 |||:|:|  
 Db 152 FLQVLMPEP 159

RESULT 11  
 ID ADK46461 standard; protein; 318 AA.

AC ADK46461;

XX 20-MAY-2004 (first entry)

DE Streptococcus pneumoniae protein, Seq ID No 2976.

XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

OS Streptococcus pneumoniae.

XX US6699703-B1.

XX 02-MAR-2004.

XX 26-MAY-2000; 2000US-00583110.

XX 02-JUL-1997; 97US-0051553P.

PR 12-MAY-1998; 98US-0085131P.

PR 30-JUN-1998; 98US-00107433.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewheat CE;  
 PI WPI; 2004-212399/20.  
 XX N-PSDB; ADK43800.  
 XX New nucleic acid molecules and polypeptides useful for diagnosing,  
 PT preventing and treating pathological conditions resulting from bacterial  
 PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
 PT screening.  
 XX Disclosure; SEQ ID NO 2976; 301pp; English.  
 XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
 CC and polypeptides. The nucleic acids and proteins are useful for  
 CC diagnosing, preventing and treating pathological conditions resulting  
 CC from bacterial infection, such as S. pneumoniae infection. These may also  
 CC be used for drug screening procedures. The present sequence represents a  
 CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
 CC data for this patent did not appear in the printed specification but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html.  
 XX Sequence 318 AA;

Query Match 77.3%; Score 34; DB 8; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLMPEP 8  
 |||:|:|  
 Db 152 FLQVLMPEP 159

RESULT 12  
 ID AAY57148 standard; protein; 2785 AA.

XX AAY57148;

XX 28-FEB-2000 (first entry)

XX Human down-regulated in metastasis (DRIM) amino acid sequence.

XX Down-regulated in metastasis; DRIM; human; antimetastatic activity;  
 KW antibody; tumour; treatment; therapy.

XX Homo sapiens.

XX WO9960116-A1.

XX 25-NOV-1999.

XX 17-MAY-1999; 99WO-EP003396.

XX 18-MAY-1998; 98EP-00303895.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX (ISIS-) ISIS INNOVATION LTD.

XX Weidle U, Tarin D;

XX WPI; 2000-053296/04.

XX N-PSDB; AAZ45136.

XX New polypeptide with antimetastatic activity, useful for therapeutic  
 PT compositions for tumor therapy.

XX Claim 2; Page 42-49; 54pp; English.

XX This is the amino acid sequence of the human down-regulated in metastasis  
 CC (DRIM) protein. The protein has antimetastatic activity. The DRIM protein  
 CC is active in both its glycosylated and unglycosylated form, and can be

CC produced by recombinant technology in prokaryotic cells. DRIM mRNA is  
 CC strongly expressed in heart, skeletal muscle, pancreas, testis and ovary  
 CC tissues. The nucleotide and protein sequences can be used to create anti-  
 CC DRIM antibodies. The nucleic acids are useful in therapeutic  
 CC compositions, especially for treating tumours. They are also useful for  
 CC activating polynucleotides from the 5' untranslated region in gene  
 CC therapy  
 CC  
 CC Sequence 2785 AA;  
 CC  
 CC Query Match 77.3%; Score 34; DB 3; Length 2785;  
 CC Best Local Similarity 77.8%; Pred. No. 1.1e+03;  
 CC Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC Qy 1 FLQLMEPV 9  
 CC |||||  
 CC Db 1056 FLDLLFEPV 1064  
 CC  
 CC RESULT 13  
 CC ABP34618  
 CC ID ABP34618 standard; protein; 117 AA.  
 CC  
 CC AC ABP34618;  
 CC  
 CC XX 08-JUL-2002 (first entry)  
 CC  
 CC XX Human ORP3591 protein, SEQ ID NO:7182.  
 CC  
 CC XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;  
 CC disease monitoring; cytokine; cell proliferation; cell differentiation;  
 CC immune modulation; haematopoiesis regulation; tissue growth;  
 CC angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
 CC thrombolytic; tumour inhibition; bodily characteristic; fertility;  
 CC behaviour; cancer; proliferative disorder; neurological disorder;  
 CC cardiovascular disease; immune system disorder; organ transplantation;  
 CC tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 CC hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
 CC vasotropic; antipsoriatic; antidiabetic; cytostatic; neutropic;  
 CC neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
 CC cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
 CC dermatological; analgesic; virucide; antibacterial; fungicide.  
 CC  
 CC OS Homo sapiens.  
 CC  
 CC XX WO200190366-A2.  
 CC  
 CC PN 29-NOV-2001.  
 CC  
 CC PD 24-MAY-2001; 2001WO-US017076.  
 CC  
 CC PF 24-MAY-2000; 2000US-0206690P.  
 CC  
 CC PR (CURA-) CURAGEN CORP.  
 CC  
 CC PA Leach MD, Shimkets RA;  
 CC  
 CC PI WPI; 2002-106200/14.  
 CC  
 CC DR N-PSDB; ABN78644.  
 CC  
 CC DR Novel human polypeptides and polynucleotides useful for diagnosing,  
 CC preventing and treating cardiovascular disease, neurodegenerative,  
 CC hyperproliferative disorders and disorders related to organ  
 CC transplantation.  
 CC  
 CC XX Claim 10; Page 2042; 2508pp; English.  
 CC  
 CC XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins  
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX

CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC polypeptides, methods of screening for modulators of ORFX expression or  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antiinfective activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration, ester  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases  
 CC  
 CC XX Sequence 117 AA;  
 CC  
 CC Qy Query Match 75.0%; Score 33; DB 5; Length 117;  
 CC Best Local Similarity 87.5%; Pred. No. 70;  
 CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC Qy 1 FLQLMEPV 8  
 CC |||||  
 CC Db 67 FLQLNEPV 74  
 CC  
 CC RESULT 14  
 CC AAG4343  
 CC ID AAG4343 standard; protein; 296 AA.  
 CC  
 CC XX AAG4343;  
 CC  
 CC AC 18-OCT-2000 (first entry)  
 CC  
 CC DT Arabidopsis thaliana protein fragment SEQ ID NO: 55533.  
 CC  
 CC DE Protein identification; signal transduction pathway; metabolic pathway;  
 CC KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 CC KW termination sequence.  
 CC  
 CC XX Arabidopsis thaliana.  
 CC  
 CC OS Arabidopsis thaliana.  
 CC  
 CC PN EP1033405-A2.  
 CC  
 CC XX 06-SEP-2000.  
 CC  
 CC PD 25-FEB-2000; 2000EP-00301439.  
 CC  
 CC PF 25-FEB-1999; 99US-0121825P.  
 CC XX 05-MAR-1999; 99US-0123180P.  
 CC PR 09-MAR-1999; 99US-0123548P.  
 CC PR 23-MAR-1999; 99US-0125788P.  
 CC PR 25-MAR-1999; 99US-0126264P.  
 CC PR 29-MAR-1999; 99US-0126785P.  
 CC PR 01-APR-1999; 99US-0127462P.  
 CC PR 06-APR-1999; 99US-0128234P.  
 CC PR 08-APR-1999; 99US-0128714P.  
 CC PR 16-APR-1999; 99US-0129845P.  
 CC PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-01351124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139482P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142220P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 28-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 12-AUG-1999; 99US-0148319P.  
PR 13-AUG-1999; 99US-0148344P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 20-AUG-1999; 99US-0149428P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.

PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 75.0%; Score 33; DB 3; Length 296;  
 Best Local Similarity 87.5%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMNEP 8  
 |||||  
 Db 217 FLOLMNEP 224

RESULT 15  
 AAG82085  
 ID AAG82085 standard; protein; 305 AA.  
 XX  
 AC AAG82085;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:1264.

KW Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;  
 endocarditis.

XX Staphylococcus epidermidis.

OS WO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US030782.

XX 09-NOV-1999; 99US-0164258P.

XX (GLAXO) GLAXO GROUP LTD.

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

XX N-PSDB; AAH52935.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
 PT useful for vaccinating against infections, e.g. endocarditis.

PS Claim 18; Page 361; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)  
 CC and (II) can have antibacterial activity and therefore can be used in  
 CC vaccination. The nucleic acids (I) may be used to produce the S.  
 CC epidermidis polypeptides (II) via the production of vectors containing

CC them which are used to produce hosts cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 CC used to vaccinate subjects and to raise antibodies against the bacteria.  
 CC The polypeptides may also be used to assay for other inhibitors of their  
 CC activity and therefore identify compounds that may be used for the  
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA  
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098  
 CC represent oligonucleotide sequences and primers which are used in the  
 CC amplification of the present invention. N.B. The present invention  
 CC specifically claims all the polynucleotide sequences given in the  
 CC sequence listing of the present specification, however the sequence  
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
 CC for SEQ ID NO:4455 to 4464

XX Sequence 305 AA;

Query Match 75.0%; Score 33; DB 4; Length 305;  
 Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMNEP 8  
 |||||  
 Db 138 FLOLMNEP 145

Search completed: December 30, 2004, 20:42:56  
 Job time : 13.4962 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:29:59 ; Search time 10.5642 Seconds  
(without alignments)  
490.180 Million cell updates/sec

Title: US-10-017-327-3  
Perfect score: 44  
Sequence: 1 FLQLMEPV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	84.1	75	2 Q8PYR9	Q8pyr9 methanosarc
2	35	79.5	47	2 Q85X02	Q85x02 pinus korai
3	35	79.5	329	2 Q8SQ02	Q8sq02 canis famil
4	35	79.5	379	2 Q03991	Q03991 daucus caro
5	34	77.3	297	1 RNHL_CAEEL	Q8u6p6 caenorhabdi
6	34	77.3	318	2 Q8DNJ4	Q8dnj4 streptococc
7	34	77.3	318	2 Q8ESS7	Q8ess7 oceanobacil
8	34	77.3	318	2 Q97NY1	Q97ny1 streptococc
9	34	77.3	471	2 Q8AVC6	Q8avc6 xenopus lae
10	34	77.3	760	2 Q9NIH0	Q9nih0 stylonychia
11	34	77.3	809	2 Q8S1S9	Q8s1s9 oryza sativ
12	34	77.3	1083	2 Q7SB63	Q7sb63 neurospora
13	34	77.3	1869	2 Q7QUP7	Q7qup7 giardia lam
14	34	77.3	2785	2 Q75691	Q75691 homo sapien
15	33	75.0	113	2 Q9SA20	Q9sa20 arabidopsis
16	33	75.0	163	2 Q85FQ9	Q85fq9 cyanidiesch
17	33	75.0	226	2 Q878Z6	Q878z6 staphylococ
18	33	75.0	252	2 Q61RD5	Q61rd5 rattus norv
19	33	75.0	252	2 AAH70960	AAH70960 rattus no
20	33	75.0	287	2 Q7NGX1	Q7ngx1 vibrio para
21	33	75.0	295	2 Q7NGE4	Q7nge4 gloeobacter
22	33	75.0	318	2 Q6GB96	Q6gb96 staphylococ
23	33	75.0	318	2 Q6GIQ7	Q6giq7 staphylococ
24	33	75.0	318	2 Q9NMF9	Q9nmf9 staphylococ
25	33	75.0	318	2 Q99VF0	Q99vp0 staphylococ
26	33	75.0	318	2 Q7A6S9	Q7a6s9 staphylococ
27	33	75.0	319	2 Q8CP29	Q8cp29 staphylococ
28	33	75.0	337	2 Q7Z5K8	Q7z5k8 homo sapien
29	33	75.0	366	2 Q84L32	Q84l32 arabidopsis
30	33	75.0	380	2 Q6TLD0	Q6tld0 brachydanio
31	33	75.0	380	2 AAQ94603	AAQ94603 brachydan

32 33 75.0 382 2 Q6PHE9 Q6phe9 brachydanio  
33 33 75.0 382 2 AAH56578 AAH56578 brachydan  
34 33 75.0 409 1 R23B\_HUMAN R23b727 homo sapien  
35 33 75.0 415 2 Q6NVC3 Q6nvc3 mus musculu  
36 33 75.0 415 2 AAH68193 AAH68193 mus muscu  
37 33 75.0 416 1 R23B\_MOUSE R23b728 mus musculu  
38 33 75.0 435 2 Q6FU70 Q6fu70 candida gla  
39 33 75.0 1569 2 Q8S012 Q8s012 oryza sativ  
40 33 75.0 2746 2 Q8IMS2 Q8ims2 drosophila  
41 33 75.0 2746 2 AAN14067 AAN14067 drosophil  
42 33 75.0 2764 2 Q01399 Q01399 drosophila  
43 33 75.0 2802 2 Q01397 Q01397 drosophila  
44 33 75.0 2802 2 Q01398 Q01398 drosophila  
45 33 75.0 2802 2 Q9VBJ2 Q9vbj2 drosophila

## ALIGNMENTS

RESULT 1  
Q8PYR9 PRELIMINARY; PRT; 75 AA.  
AC Q8PYR9  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein MM0791.  
GN OrderedLocusNames=MM0791;  
OS Methanosarcina mazel (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Goel / Goel / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=42120827; PubMed=12125824;  
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
RA Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,  
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,  
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RT "The genome of Methanosarcina mazel: evidence for lateral gene  
transfer between Bacteria and Archaea.";  
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
DR EMBL; AE013304; AAM30487.1; -.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 75 AA; 9035 MW; 695543A5C06F505D CRC64;

Query Match 84.1%; Score 37; DB 2; Length 75;  
Best Local Similarity 87.5%; Pred. No. 4.5;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMEPV 8  
|||  
Db 40 FLQLLEPV 47

RESULT 2  
Q85X02 PRELIMINARY; PRT; 47 AA.  
AC Q85X02  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ORF470.  
OS Pinus koraiensis (Korean pine).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Pinaceae; Pinus.  
OX NCBI\_TaxID=88728;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Noh E.W., Lee J.S., Choi Y.I., Han M.S., Yi Y.S., Han S.U.;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY228468; AA074065.1; -.  
DR GO; GO:0009507; C:chloroplast; IEA.  
KW Chloroplast.  
SQ SEQUENCE 47 AA; 5510 MW; B72C7950F3876DCD CRC64;  
  
Query Match 79.5%; Score 35; DB 2; Length 47;  
Best Local Similarity 87.5%; Pred. No. 7.7;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 FLQLMEPV 8  
Db 25 FQLLFEP 32  
  
RESULT 3  
Q8SQ02 PRELIMINARY; PRT; 329 AA.  
ID Q8SQ02;  
AC Q8SQ02;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Pregnane X receptor (Fragment).  
GN Name=PX;  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21977871; PubMed=11981033;  
RA Moore L.B., Maglich J.M., McKee D.D., Wisely B., Willson T.M.,  
RA Klier S.A., Lambert M.H., Moore J.T.;  
RT "Pregnane X receptor (PXR), constitutive androstane receptor (CAR),  
RT and benzate X receptor (BXR) define three pharmacologically distinct  
RT classes of nuclear receptors.";  
RL Mol. Endocrinol. 16:977-986(2002).  
DR EMBL; AF454670; AAM10632.1; -.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; P:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR00536; Hrmn\_recept\_lig.  
DR InterPro; IPR01723; Stchrnm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR00324; Vitd\_receptor.  
DR Pfam; PF01104; Hormone\_recep. 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00350; VITAMINDR.  
DR SMART; SM00430; HOLH; 1.  
KW Receptor.  
FT NON TER.  
SQ SEQUENCE 329 AA; 37479 MW; 237F59C90C367474 CRC64;  
  
Query Match 79.5%; Score 35; DB 2; Length 329;  
Best Local Similarity 77.8%; Pred. No. 59;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 FLQLMEPV 9  
Db 210 FQLLFEPV 218  
  
RESULT 4  
O03991 PRELIMINARY; PRT; 379 AA.  
ID O03991;  
AC O03991;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE RAD23 protein, isoform II.  
OS Daucus carota (Carrot).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asteride;  
OC campanulids; Apiales; Apiaceae; Scandiceae; Daucinae;  
OC Daucus.  
OX NCBI\_TaxID=4039;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98345997; PubMed=9681019;  
RA Sturm A., Leinhardt S.;  
RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in  
RT yeast.";  
RL Plant J. 13:815-821(1998).  
DR EMBL; Y12014; CAA72742.1; -.  
DR PIR; T14337; T14337.  
DR HSSP; P54725; 1F41.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.  
DR InterPro; IPR004806; Rad23.  
DR InterPro; IPR006636; STI1.  
DR InterPro; IPR000449; UBA.  
DR InterPro; IPR009060; UBA-like.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00627; UBA; 2.  
DR Pfam; PF00240; ubiquitin; 1.  
DR PRINTS; PR01839; RAD23PROTEIN.  
DR SMART; SM00727; STI1; 1.  
DR SMART; SM00165; UBA; 2.  
DR SMART; SM00213; UBQ; 1.  
DR TIGRFAMs; TIGR00601; rad23; 1.  
DR PROSITE; PS50030; UBA; 2.  
DR PROSITE; PS50053; UBIQUITIN 2; 1.  
SQ SEQUENCE 379 AA; 40530 MW; B26697B439CC5929 CRC64;  
  
Query Match 79.5%; Score 35; DB 2; Length 379;  
Best Local Similarity 77.8%; Pred. No. 68;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 FLQLMEPV 9  
Db 302 FQLLINEPV 310  
  
RESULT 5  
ID RNHL CAEEL STANDARD; PRT; 297 AA.  
AC Q9U6P6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ribonuclease H1 large subunit (EC 3.1.26.-) (RNase H1 large subunit)  
DE (RNase H(35)).  
GN Name=rnh-2; Synonyms=rnh2; ORFNames=T13H5.7;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22300325; PubMed=12411600;  
RA Aruchandran A.P., Cerritelli S.M., Bowen N.J., Chen X., Krause M.W.,  
RA Crouch R.J.;  
RT "Multiple ribonuclease H-encoding genes in the Caenorhabditis elegans  
RT genome contrasts with the two typical ribonuclease H-encoding genes in  
RT the human genome.";  
RL Mol. Biol. Evol. 19:1910-1919(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Lightning J.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Degrades the ribonucleotide moiety on RNA-DNA hybrid  
CC molecules. Participates in DNA replication (By similarity).  
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
CC phosphomononucleoside.



CC -!- SIMILARITY: Belongs to the RNase HII family. Eukaryotic subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----

DR EMBL: AF181619; AAF01208.1; --  
 DR EMBL: Z66524; CAC70103.1; --  
 DR HSRF: Q57599; IEKE.  
 DR WormPep: T13H5.7; CR28960.  
 DR InterPro: IPR001352; RNase HII/HIII.  
 DR InterPro: IPR004649; RnhII-  
 DR Pfam: PF01351; RNase HII; 1.  
 DR TIGRFAMs: TIGR00729; RnhII; 1.  
 KW Endonuclease; Hydrolase; Nuclease.  
 FT ACT\_SITE 27 27 By similarity.  
 FT ACT\_SITE 138 138 By similarity.  
 FT ACT\_SITE 166 166 By similarity.  
 SQ SEQUENCE 297 AA; 33191 MW; 233C11EDD4A7B5B2 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 297;  
 Best Local Similarity 77.8%; Pred. No. 87;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLMPEV 9

Db 217 FLQLSVEPV 225

RESULT 6

Q8DNJ4 ID Q8DNJ4 PRELIMINARY; PRT; 318 AA.  
 AC Q8DNJ4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE ABC transporter membrane-spanning permease-ferric iron transport.  
 GN Name=fatC; OrderedLocNames=spr1685;  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=171101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-255 / R6;  
 RX MEDLINE=21429245; PubMed=11544234;  
 RA Hopkins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,  
 RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,  
 RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,  
 RA McHarris S.M., McHennery M., McLeaster K., Mundy C.W., Niclas T.I.,  
 RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,  
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,  
 RA Glass J.I.;  
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";  
 RL J. Bacteriol. 183:5709-5717(2001).  
 DR EMBL: AE008534; AAL00488.1; --  
 DR PIR: C98082.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005215; P:transporter activity; IEA.  
 DR GO: GO:0006810; P:transport; IEA.  
 DR InterPro: IPR000522; FecD.  
 DR Pfam: PF01032; FecCD; 1.  
 KW Complete proteome.

SQ SEQUENCE 318 AA; 36296 MW; 9C1894E4F57C3210 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 94;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FLQLMPE 8  
 Db 152 FLQVLMDP 159

RESULT 7

Q8ESS7 ID Q8ESS7 PRELIMINARY; PRT; 318 AA.  
 AC Q8ESS7;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ferrichrome ABC transporter permease.  
 GN OrderedLocNames=OB0540;  
 OS Oceanobacillus iheyensis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
 OX NCBI\_TaxID=182710;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HTE831;  
 RX MEDLINE=22220767; PubMed=12235376;  
 RA Takami H., Takaki Y., Uchiyama I.;  
 RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya  
 RT Ridge and its unexpected adaptive capabilities to extreme  
 RT environments.";  
 RL Nucleic Acids Res. 30:3927-3935(2002).  
 DR EMBL: AP004594; BAC12496.1; --  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005215; P:transporter activity; IEA.  
 DR GO: GO:0006810; P:transport; IEA.  
 DR InterPro: IPR000522; FecD.  
 DR Pfam: PF01032; FecCD; 1.  
 KW Complete proteome.

SQ SEQUENCE 318 AA; 35775 MW; 3A53578580A52E71 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 318;  
 Best Local Similarity 66.7%; Pred. No. 94;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLMPEV 9

Db 150 FLQVLIDPV 158

RESULT 8

Q97NY1 ID Q97NY1 PRELIMINARY; PRT; 318 AA.  
 AC Q97NY1;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Iron-compound ABC transporter, permease protein.  
 GN OrderedLocNames=SP1870;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-334 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 RA Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,  
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 RT pneumoniae.";

```

RL Science 293:498-506(2001).
DR EMBL; AE007478; AAK75942.1; -.
DR PIR; S95218; E95218.
DR TIGR; SP1870; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FcCD.
DR Pfam; PF01032; FcCD; 1.
KW Complete proteome.
SQ SEQUENCE 318 AA; 36234 MW; 72BB0DC3DC0BFD13 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 318;
Best Local Similarity 75.0%; Pred. NO. 94;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLQLLMEPV 8
Db 152 FLQVLMDF 159

RESULT 9
Q8AVC6 Q8AVC6 PRELIMINARY; PRT; 471 AA.
AC Q8AVC6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 23, Last annotation update)
DE MG52635 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schermer A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative".
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC041717; AAH41717.1; -.

```

```

DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS00850; MFS; 1.
SQ SEQUENCE 471 AA; 51058 MW; 1504C5B99AA70B01 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 471;
Best Local Similarity 66.7%; Pred. NO. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLQLLMEPV 9
Db 121 FLQVLMDF 129

RESULT 10
Q9NIH0 Q9NIH0 PRELIMINARY; PRT; 760 AA.
AC Q9NIH0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Elongation factor 2 (Fragment).
GN Name=EF2;
OS Stylonychia mytilus.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
OC Stichotrichida; Oxytrichidae; Stylonychia.
OX NCBI_TaxID=5950;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20269356; PubMed=10811219;
RA Moreira D., Le Guyader H., Philippe H.;
RT "The origin of red algae and the evolution of chloroplasts."
RL Nature 405:69-72(2000).
DR EMBL; AF213664; AAF71707.1; -.
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0008411; P:protein biosynthesis; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR009022; EFG III V.
DR InterPro; IPR005517; EFG IV.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00579; EFG C; 1.
DR Pfam; PF03764; EFG IV; 1.
DR Pfam; PF00009; GTP_EFTU_D2; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT
DR TIGRPFams; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
FT NON_TER 1
FT NON_TER 760
SQ SEQUENCE 760 AA; 84711 MW; 6CD14EE3F1E6590D CRC64;

Query Match 77.3%; Score 34; DB 2; Length 760;
Best Local Similarity 55.6%; Pred. NO. 2.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLQVLMDF 9
Db 237 FVQVLMDF 245

RESULT 11
Q8S159 Q8S159 PRELIMINARY; PRT; 809 AA.
ID Q8S159
AC Q8S159;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE P0042A10.6 protein.  
 GN Name=P0042A10.6;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arica K., Hamada M., Harada C.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikano M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Teraawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo I., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RL "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 CC -!- SIMILARITY: Contains 7 WD repeats.  
 CC EMBL; AP003343; BAB90064.1; -;  
 DR HSBF; P16849; IERJ.  
 DR Gramene; Q8S159; -;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004672; F:protein kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006458; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001680; WD40.  
 DR InterPro; IPR011046; WD40\_like.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINERPT.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000018; WD40; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE; PS00082; WD\_REPEATS\_2; 2.  
 DR PROSITE; PS02094; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 809 AA; 88909 MW; B1F770AD98B36A4A CRC64;  
 Query Match 77.3%; Score 34; DB 2; Length 809;  
 Best Local Similarity 77.8%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 FQLLMPEPV 9  
 Db 237 FCQLLMPEPV 245  
 RESULT 12  
 Q7SB63 PRELIMINARY; PRT; 1083 AA.  
 ID Q7SB63  
 AC Q7SB63;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DE Hypothetical protein.  
 GN Name=NCU06278.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Renman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Tanakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kanal M., Kamysseis M., Maucelli E., Bielek C., Rudd S., Frisman D.,  
 RA Krystofova S., Rasmussen C., Metzberg R.I., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,  
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Navig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nuebaum C., Birren B.;  
 RL "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0:0-0(2003).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABX01000173; EAA33646.1; -;  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0003746; P:translation elongation factor activity; IEA.  
 DR GO; GO:0006412; P:protein biosynthesis; IEA.  
 DR GO; GO:0006414; P:translational elongation; IEA.  
 DR InterPro; IPR000640; ERG\_C.  
 DR InterPro; IPR005517; ERG\_IV.  
 DR InterPro; IPR004161; EFTU\_D2.  
 DR InterPro; IPR000795; ProtSyn\_GTPbind.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00679; ERG\_C; 1.  
 DR Pfam; PF03764; ERG\_IV; 1.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR TIGRfams; TIGR00231; small\_GTP; 1.  
 KW GTP-binding; Hypothetical protein; Protein biosynthesis.  
 SQ SEQUENCE 1083 AA; 11919 MW; A2CD42DD89C64731 CRC64;  
 Query Match 77.3%; Score 34; DB 2; Length 1083;  
 Best Local Similarity 55.6%; Pred. No. 3.4e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FQLLMPEPV 9  
 Db 312 FVQLVLEPI 320  
 RESULT 13  
 Q7QUP7 PRELIMINARY; PRT; 1869 AA.  
 ID Q7QUP7  
 AC Q7QUP7;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DE GUP\_47\_16110.21719.  
 OS Giardia lamblia ATCC 50803.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.  
 OX NCBI\_TaxID=184922;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
 RA Olsen G.J., Sogin M.L.;  
 RT "Draft sequence of the Giardia lamblia genome.";  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AACB01000091; EAA38765.1; -;

SQ SEQUENCE 1869 AA; 209537 MW; 55EE7149EDA47640 CRC64;  
Query Match 77.3%; Score 34; DB 2; Length 1869;  
Best Local Similarity 75.0%; Pred. No. 6e+02; 0; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 0;  
QY 1 FLOLLMEPV 8  
Db 1481 FLELLLEP 1488  
RESULT 14  
O75691  
ID O75691 PRELIMINARY; PRT; 2785 AA.  
AC O75691;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE DRIM protein.  
GN Name=drim;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98338066; PubMed=9673349;  
RA Schwirzke M., Gairke A., Bork P., Tarin D., Weidle U.H.;  
RT "Differential gene expression in mammary carcinoma cell lines:  
identification of DRIM, a new gene down-regulated in metastasis.";  
RL Anticancer Res. 18:1409-1421(1998).  
RL EMBL; AJ006778; CAA07243.1; -.  
DR SWISS-2DPAGE; O75691; -.  
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR011430; DRIM.  
DR Pfam; PF07539; DRIM; 1.  
SQ SEQUENCE 2785 AA; 318423 MW; 6A2BBE9E82ADB983 CRC64;  
Query Match 77.3%; Score 34; DB 2; Length 2785;  
Best Local Similarity 77.8%; Pred. No. 9.1e+02; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 2;  
QY 1 FLOLLMEPV 9  
Db 1056 FLOLLLEPV 1064  
RESULT 15  
Q9SA20  
ID Q9SA20 PRELIMINARY; PRT; 113 AA.  
AC Q9SA20;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE F309.1 protein.  
GN Name=F309.1;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Vysotskaia V.S., Schwartz J., Yu G., Toriumi M., Lenz C., Liu S.,  
RA Lee J., Li J., Kremenetskaia I., Liu A., Luros J., Gonzalez A.,  
RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,  
RA Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,  
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Theologis;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006341; AAD34676.1; -.  
DR PIR; H86296; H86296.  
DR HSP; P54725; 1F4I.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.  
DR InterPro; IPR009020; Prot\_inh\_propept.  
DR InterPro; IPR004806; Rad23.  
DR InterPro; IPR000449; UBA.  
DR InterPro; IPR009060; UBA\_like.  
DR Pfam; PF00627; UBA; 1.  
DR PRINTS; PR01839; RAD23PROTEIN.  
DR SMART; SM00165; UBA; 1.  
DR PROSITE; PSS0030; UBA; 1.  
SQ SEQUENCE 113 AA; 12947 MW; 8BE7CB15C9E12380 CRC64;  
Query Match 75.0%; Score 33; DB 2; Length 113;  
Best Local Similarity 87.5%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLOLLMEPV 8  
Db 34 FLOLLLEPV 41  
Search completed: December 30, 2004, 20:50:49  
Job time : 13.5642 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:34:10 ; Search time 2.06297 Seconds  
(without alignments)  
289.321 Million cell updates/sec

Title: US-10-017-327-5

Perfect score: 44

Sequence: 1 FLQLEFDAV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	75.0	420	4	US-09-248-796A-18548
2	32	72.7	419	4	US-09-252-991A-27064
3	32	72.7	523	4	US-09-270-767-43156
4	32	72.7	633	4	US-09-919-060-13
5	32	72.7	2289	3	US-09-051-019-2
6	31	70.5	210	1	US-07-667-276A-2
7	31	70.5	287	4	US-09-543-681A-7259
8	31	70.5	379	4	US-09-270-767-32891
9	31	70.5	379	4	US-09-270-767-48108
10	31	70.5	524	4	US-09-252-991A-17710
11	31	70.5	803	4	US-09-489-039A-12742
12	30	68.2	191	4	US-09-621-976-6999
13	30	68.2	299	4	US-09-584-568C-6
14	30	68.2	300	4	US-09-194-146-6
15	30	68.2	333	4	US-09-248-796A-17292
16	30	68.2	370	3	US-09-134-001C-3403
17	30	68.2	370	4	US-09-710-279-696
18	30	68.2	370	4	US-09-710-279-1328
19	30	68.2	582	3	US-09-518-914-21
20	30	68.2	1176	4	US-09-976-594-793
21	29	65.9	57	4	US-09-513-999C-7337
22	29	65.9	130	3	US-09-302-769-29
23	29	65.9	134	4	US-09-270-767-60072
24	29	65.9	141	4	US-09-621-976-6981
25	29	65.9	141	4	US-09-621-976-6982
26	29	65.9	141	4	US-09-513-999C-7895
27	29	65.9	178	4	US-09-153-447-21

#### ALIGNMENTS

##### RESULT 1

US-09-248-796A-18548  
; Sequence 18548, Application US/09248796A  
; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 18548

; LENGTH: 420

; TYPE: PRT

; ORGANISM: Candida albicans

; US-09-248-796A-18548

Query Match 75.0%; Score 33; DB 4; Length 420;  
Best Local Similarity 85.7%; Pred. No. 78;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLEFDAV 9

Db 339 QLEFDAL 345

##### RESULT 2

US-09-252-991A-27064

; Sequence 27064, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27064

; LENGTH: 419

; TYPE: PRT

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27064

Query Match          72.7%; Score 32; DB 4; Length 419;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLEFDV 9
Db 226 LQVEFDAL 233

RESULT 3
US-09-270-767-43156
; Sequence 43156, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1993-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43156
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43156

Query Match          72.7%; Score 32; DB 4; Length 523;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLEFDA 8
Db 477 FLQLQFNA 484

RESULT 4
US-09-919-060-13
; Sequence 13, Application US/09919060
; Patent No. 6638744
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES
; FILE REFERENCE: AD-1
; CURRENT APPLICATION NUMBER: US/09/919,060
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/224,486
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-919-060-13

Query Match          72.7%; Score 32; DB 4; Length 633;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFD 7
Db 389 FQLEKFD 395

RESULT 5
US-09-051-019-2

```

```

; Sequence 2, Application US/09051019
; Patent No. 6103229
; GENERAL INFORMATION:
; APPLICANT: KAHMANN, Regine and QUADBECK-SEEGER, Claudia
; TITLE OF INVENTION: Regulatory gene from Ustilago maydis
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinhaus
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage
; COMPUTER: IBM AT-compatible, Pentium processor
; OPERATING SYSTEM: Windows 98
; SOFTWARE: WordPerfect version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051,019
; FILING DATE: 31-MAR-1998
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-051-019-2

Query Match          72.7%; Score 32; DB 3; Length 2289;
Best Local Similarity 66.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQLEFDV 9
Db 1510 FLPLEWDAL 1518

RESULT 6
US-07-667-276A-2
; Sequence 2, Application US/07667276A
; Patent No. 5470971
; GENERAL INFORMATION:
; APPLICANT: Kondo, Keiji
; APPLICANT: Inouye, Masayori
; TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND
; TITLE OF INVENTION: APPLICATIONS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 S. Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/667,276A
; FILING DATE: 11-MAR-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377.5351P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 210 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-667-276A-2

Query Match 70.5%; Score 31; DB 1; Length 210;  
Best Local Similarity 55.6%; Pred. No. 97;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQLEFDV 9  
| : : : :  
Db 78 FSELDVDAI 86

## RESULT 7

US-09-543-681A-7259  
; Sequence 7259, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 7259  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-7259

Query Match 70.5%; Score 31; DB 4; Length 287;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLEFDV 9  
| : : : :  
Db 139 QVEFDV 145

## RESULT 8

US-09-270-767-32891  
; Sequence 32891, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 32891  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-32891

Query Match 70.5%; Score 31; DB 4; Length 379;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQLEFDV 9  
| : : : :  
Db 360 FLSMQYDAV 368

## RESULT 9

US-09-270-767-48108  
; Sequence 48108, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48108  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-48108

Query Match 70.5%; Score 31; DB 4; Length 379;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQLEFDV 9  
| : : : :  
Db 360 FLSMQYDAV 368

## RESULT 10

US-09-252-991A-17710  
; Sequence 17710, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17710  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17710

Query Match 70.5%; Score 31; DB 4; Length 524;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQLEFDV 9  
| : : : :  
Db 245 FLQVEVDAL 253

## RESULT 11

US-09-489-039A-12742  
; Sequence 12742, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12742  
; LENGTH: 803



```
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12742

Query Match      70.5%; Score 31; DB 4; Length 803;
Best Local Similarity 77.8%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLQLEFDV 9
      |||||
Db      577 FLHLLFDV 585

RESULT 12
US-09-621-976-6999
; Sequence 6999, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6999
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6999

Query Match      68.2%; Score 30; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LQLEFD 7
      |||||
Db      25 LQLEFD 30

RESULT 13
US-09-584-568C-6
; Sequence 6, Application US/09584568C
; Patent No. 6500657
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria, Alexandra et al.
; TITLE OF INVENTION: 33167, A NOVEL HUMAN HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MNI-140
; CURRENT APPLICATION NUMBER: US/09/584.568C
; CURRENT FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/193,954
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-584-568C-6

Query Match      68.2%; Score 30; DB 4; Length 299;
Best Local Similarity 55.6%; Pred. No. 2.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLQLEFDV 9
      |||||
Db      266 FPQVQFDV 274

RESULT 14
US-09-194-146-6
```

```
; Sequence 6, Application US/09194146
; Patent No. 6458557
; GENERAL INFORMATION:
; APPLICANT: Miller, Brian
; APPLICANT: Diaz-Torres, Maria
; TITLE OF INVENTION: Gram-Positive Microorganism Formate Pathway
; FILE REFERENCE: GC395-US
; CURRENT APPLICATION NUMBER: US/09/194.146
; CURRENT FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: GB 9724627.6
; PRIOR FILING DATE: 1997-11-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Bacillus
US-09-194-146-6
```

```
Query Match      68.2%; Score 30; DB 4; Length 300;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FLQLEFD 7
      |||||
Db      63 FLRIEFD 69
```

```
RESULT 15
US-09-248-796A-17292
; Sequence 17292, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
```

```
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074.725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096.409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17292
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17292
```

```
Query Match      68.2%; Score 30; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FLQLEF 6
      |||||
Db      129 FLQLEF 134
```

```
Search completed: December 30, 2004, 20:53:55
Job time : 3.06297 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:35:11 ; Search time 8.81864 Seconds  
(without alignments)  
367.126 Million cell updates/sec

Title: US-10-017-327-5  
Perfect score: 44  
Sequence: 1 FLQLEFDV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	9	11 US-09-870-216C-5	Sequence 5, Appli
2	44	100.0	9	13 US-10-017-327-5	Sequence 5, Appli
3	35	79.5	219	17 US-10-425-115-212999	Sequence 212999,
4	35	79.5	461	17 US-10-425-115-213000	Sequence 213000,
5	35	79.5	654	14 US-10-369-493-9433	Sequence 9433, Ap
6	35	79.5	680	14 US-10-369-493-17753	Sequence 17753, A
7	34	77.3	156	16 US-10-437-963-198066	Sequence 198066,
8	33	75.0	125	15 US-10-424-599-274136	Sequence 274136,
9	33	75.0	388	14 US-10-369-493-12756	Sequence 12756, A
10	32	72.7	225	15 US-10-282-122A-49767	Sequence 49767, A
11	32	72.7	274	15 US-10-282-122A-51999	Sequence 51999, A
12	32	72.7	414	14 US-10-260-937-5	Sequence 5, Appli
13	32	72.7	414	14 US-10-260-937-49	Sequence 49, Appli

14	32	72.7	414	14	US-10-260-937-59	Sequence 59, Appli
15	32	72.7	551	15	US-10-382-248-8	Sequence 8, Appli
16	32	72.7	551	17	US-10-741-853A-2	Sequence 2, Appli
17	32	72.7	562	16	US-10-741-601-515	Sequence 515, App
18	32	72.7	580	14	US-10-260-937-34	Sequence 34, Appli
19	32	72.7	599	10	US-09-953-067A-3	Sequence 3, Appli
20	32	72.7	599	14	US-10-097-340-264	Sequence 264, App
21	32	72.7	599	14	US-10-260-937-47	Sequence 47, Appli
22	32	72.7	599	15	US-10-382-248-6	Sequence 6, Appli
23	32	72.7	599	16	US-10-741-601-514	Sequence 514, App
24	32	72.7	600	14	US-10-260-937-48	Sequence 48, Appli
25	32	72.7	602	10	US-09-953-067A-7	Sequence 7, Appli
26	32	72.7	602	14	US-10-260-937-51	Sequence 51, Appli
27	32	72.7	602	14	US-10-260-937-52	Sequence 52, Appli
28	32	72.7	603	14	US-10-260-937-58	Sequence 58, Appli
29	32	72.7	606	14	US-10-260-937-50	Sequence 50, Appli
30	32	72.7	629	14	US-10-260-937-15	Sequence 15, Appli
31	32	72.7	630	17	US-10-783-297A-9	Sequence 9, Appli
32	32	72.7	630	17	US-10-783-297A-11	Sequence 11, Appli
33	32	72.7	632	14	US-10-260-937-60	Sequence 13, Appli
34	32	72.7	633	9	US-09-919-060-13	Sequence 2, Appli
35	32	72.7	633	14	US-10-260-937-2	Sequence 13, Appli
36	32	72.7	633	15	US-10-679-140-13	Sequence 8814, Ap
37	32	72.7	641	14	US-10-369-493-8814	Sequence 17121, A
38	32	72.7	756	14	US-10-369-493-17121	Sequence 160478,
39	31	70.5	141	16	US-10-437-963-160478	Sequence 194355,
40	31	70.5	167	16	US-10-437-963-194355	Sequence 4, Appli
41	31	70.5	188	14	US-10-308-460-4	Sequence 2, Appli
42	31	70.5	190	14	US-10-308-460-2	Sequence 6, Appli
43	31	70.5	190	14	US-10-308-460-6	Sequence 2688, Ap
44	31	70.5	195	15	US-10-276-774-2688	Sequence 194570,
45	31	70.5	212	16	US-10-437-963-194570	

#### ALIGNMENTS

RESULT 1  
US-09-870-216C-5  
; Sequence 5, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; ORGANISM: US-09-870-216C-5

Query Match 100.0%; Score 44; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLEFDV 9  
| | | | |  
Db 1 FLQLEFDV 9

RESULT 2  
US-10-017-327-5  
; Sequence 5, Application US/10017327

```
; Publication NO. US20020155471A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
; FILE REFERENCE: G2 2101.20
; CURRENT APPLICATION NUMBER: US/10/017,327
; PRIORITY FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-327-5

Query Match      100.0%; Score 44; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLEFQDAV 9
DB 1 FLQLEFQDAV 9

RESULT 3
US-10-425-115-212999
; Sequence 212999, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 212999
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_125856C.1.pep
US-10-425-115-212999

Query Match      79.5%; Score 35; DB 17; Length 219;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFQDAV 9
DB 72 FLQLEIDSV 80

RESULT 4
US-10-425-115-213000
; Sequence 213000, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 213000
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_125856C.1.pep
US-10-425-115-213000

Query Match      79.5%; Score 35; DB 17; Length 219;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFQDAV 9
DB 72 FLQLEIDSV 80

RESULT 5
US-10-369-493-9433
; Sequence 9433, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIORITY FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIORITY FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9433
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9433

Query Match      79.5%; Score 35; DB 14; Length 654;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFQDAV 9
DB 89 FLQIEHDAV 97

RESULT 6
US-10-369-493-17753
; Sequence 17753, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIORITY FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIORITY FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17753
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_125857C.1.pep
US-10-425-115-213000
```

```
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-17753

Query Match          79.5%; Score 35; DB 14; Length 680;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFDV 9
   |||: |||
Db 114 FLQIEHDAV 122

RESULT 7
US-10-437-963-198066
; Sequence 198066, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198066
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93763C.1.pap
US-10-437-963-198066

Query Match          77.3%; Score 34; DB 16; Length 156;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLEFDV 9
   |||||
Db 78 QLEFDV 84

RESULT 8
US-10-424-599-274136
; Sequence 274136, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 274136
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(125)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_89567C.1.pap
US-10-424-599-274136

Query Match          75.0%; Score 33; DB 15; Length 125;
Best Local Similarity 77.8%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLQLEFDV 9
   |||: |||
Db 77 FLQLPFDAV 85

RESULT 9
US-10-369-493-12756
; Sequence 12756, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12756
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(388)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12756

Query Match          75.0%; Score 33; DB 14; Length 388;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEFDV 9
   |||: |||
Db 54 LQLEFDV 61

RESULT 10
US-10-282-122A-49767
; Sequence 49767, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 49767
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-282-122A-49767

Query Match          72.7%; Score 32; DB 15; Length 225;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLOLEFDPV 9
DB      111 YLKLEFDPV 119

RESULT 11
US-10-282-122A-51999
; Sequence 51999, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636

```

```

; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 51999
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51999

Query Match          72.7%; Score 32; DB 15; Length 274;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLOLEFDPV 9
DB      205 FLKLEFDPV 213

RESULT 12
US-10-260-937-5
; Sequence 5, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Daniel
; APPLICANT: Chandrasekharan, N. Vishvanath
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 07913-007001
; CURRENT APPLICATION NUMBER: US/10/260,937
; CURRENT FILING DATE: 2002-09-28
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Canis familiaris
US-10-260-937-5

Query Match          72.7%; Score 32; DB 14; Length 414;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLOLEFDPV 7
DB      170 FLQLEFDPV 176

RESULT 13
US-10-260-937-49
; Sequence 49, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Daniel
; APPLICANT: Chandrasekharan, N. Vishvanath
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 07913-007001
; CURRENT APPLICATION NUMBER: US/10/260,937
; CURRENT FILING DATE: 2002-09-28
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225

```

```
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Canis familiaris
US-10-260-937-49
```

```
Query Match 72.7%; Score 32; DB 14; Length 414;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FQLEFD 7
   ||||:|
Db 170 FLQKFD 176
```

```
RESULT 14
US-10-260-937-59
; Sequence 59, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Daniel
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 07913-007001
; CURRENT APPLICATION NUMBER: US/10/260,937
; CURRENT FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Canis familiaris
US-10-260-937-59
```

```
Query Match 72.7%; Score 32; DB 14; Length 414;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FQLEFD 7
   ||||:|
Db 170 FLQKFD 176
```

```
RESULT 15
US-10-382-248-8
; Sequence 8, Application US/10382248
; Publication No. US20040058347A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook, et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-568C
; CURRENT APPLICATION NUMBER: US/10/382,248
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 60/361,974
; PRIOR FILING DATE: 2002-03-06
```

```
; PRIOR APPLICATION NUMBER: 60/365,477
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/401,661
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 8
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-382-248-8
```

```
Query Match 72.7%; Score 32; DB 15; Length 551;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FQLEFD 7
   ||||:|
Db 307 FLQKFD 313
```

```
Search completed: December 30, 2004, 21:00:32
Job time : 9.81864 secs
```

**This Page Blank (uspio)**





XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen eIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents synthetic epitope 2 of human cancer antigen eIF3  
 XX Sequence 9 AA;

Query Match 100.0%; Score 44; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLEFDDAV 9  
 Db |||||  
 1 FLQLEFDDAV 9

RESULT 2  
 ABR82214  
 ID ABR82214 standard; peptide; 9 AA.

AC ABR82214;  
 DT 13-OCT-2003 (first entry)  
 XX Human antigen eIF3 derived compound 2.  
 DE Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
 KW cytosolic; gene therapy; human; antigen.

XX Synthetic.  
 OS Homo sapiens.  
 XX WO2003050543-A1.  
 XX 19-JUN-2003.  
 PD 05-DEC-2001; 2001WO-US047997.  
 XX 05-DEC-2001; 2001WO-US047997.  
 XX (GENZ ) GENZYME CORP.  
 XX Nicolette CA;  
 PI WPI; 2003-532936/50.  
 DR N-PSDB; ACC85031.  
 XX Aiding in the diagnosis of a neoplastic condition, useful for treating  
 PT cancer and related malignancies comprises determining the amount of  
 PT expression of an eIF3 protein in a test sample isolated from the cell or  
 PT tissue.  
 XX Claim 12; Page 30; 77pp; English.

XX The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic

CC condition or susceptibility to a neoplastic condition based on the amount  
 CC of expression of the eIF3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. Sequences ABR82213-16  
 CC represent compounds derived from the human antigen eIF3  
 XX Sequence 9 AA;

Query Match 100.0%; Score 44; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLEFDDAV 9  
 Db |||||  
 1 FLQLEFDDAV 9

RESULT 3  
 ABB08367

ID ABB08367 standard; protein; 352 AA.

AC ABB08367;  
 XX 07-MAY-2002 (first entry)  
 DT Human cancer antigen eIF3 variant 2 amino acid sequence.  
 DE Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 242 /note= "wild-type Asn is replaced by Phe"  
 FT Domain 242 /note= "HLA-2 binding residue"  
 FT Domain 243 /note= "HLA-2 binding residue"  
 FT Domain 244.249 /note= "T-cell receptor (TCR) binding domain"  
 FT Misc-difference 246 /note= "wild-type Leu is replaced by Glu"  
 FT Misc-difference 247 /note= "wild-type Met is replaced by Phe"  
 FT Misc-difference 249 /note= "wild-type Arg is replaced by Ala"  
 FT Domain 250 /note= "HLA-2 binding residue"

XX WO200192307-A2.

PN 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017456.

XX 31-MAY-2000; 2000US-0209391P.

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

XX (GENZ ) GENZYME CORP.

PA Nicolette CA;

XX WPI; 2002-139606/18.

XX New therapeutic compounds useful against human ovarian cancer, for  
 PT modulating immune response in a subject, and for generating antibodies  
 PT that specifically recognize and bind to these molecules.

XX Claim 7; Page; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen E1F3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies. The polynucleotides can be used as primers for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human cancer antigen e1F3 variant 2 amino acid sequence.  
 CC Note: This sequence is not present in the specification, but may be  
 CC created from the sequence of the wild-type human cancer antigen e1F3  
 CC sequence given in ABB08360  
 XX  
 SQ Sequence 352 AA;

Query Match 100.0%; Score 44; DB 5; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLEFDV 9  
 DB 242 FLQLEFDV 250  
 |||||

RESULT 4  
 AAB62030  
 ID AAB62030 standard; protein; 1278 AA.  
 XX  
 AC AAB62030;  
 XX  
 DT 14-MAY-2001 (first entry)  
 XX  
 DE Recombinant P. furiosus helicase 7.  
 XX  
 KW MCM; minichromosome maintenance protein; archaeal polypeptide; PCNA;  
 KW RFC-P38; RFC-P55; RFA; CDC6; FEN-1; dUTPase; ligase; helicase dna2; PCR;  
 KW helicase 7; nucleic acid amplification; polymerase chain reaction.  
 XX  
 OS Pyrococcus furiosus.  
 XX  
 PF WO200109347-A2.  
 XX  
 PD 08-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000WO-US020532.  
 XX  
 PR 30-JUL-1999; 99US-0146580P.  
 XX  
 PA (STRA-) STRATAGENE.  
 XX  
 PI Hogrefe HH, Cline JM, Hansen CJ, Borns MC;  
 XX  
 DR WPI; 2001-182959/18.  
 DR N-PSDB; AAF57034.  
 XX  
 PT Composition for improving nucleic acid polymerase reactions, useful e.g.  
 PT in synthesis or amplification, contains at least one archaeal accessory  
 PT protein.  
 XX  
 PS Claim 182; Fig 34; 147pp; English.  
 XX  
 CC The invention provides a composition (A) for enhancing nucleic acid  
 CC polymerase reactions that comprises an archaeal MCM (minichromosome  
 CC maintenance protein) and at least one of the archaeal polypeptides (PCNA,

CC RFC-P38 or -P55, RFA, CDC6, FEN-1, dUTPase, ligase, helicase dna2, or  
 CC helicases 2-8). (A) And similar compositions containing different  
 CC combinations of accessory proteins, are used to improve performance of  
 CC synthesis, amplification, mutagenizing, labeling and detecting reactions,  
 CC e.g. for gene characterization, cloning, detection of allelic variants,  
 CC diagnosis and screening for disease, particularly where done by  
 CC polymerase chain reaction (PCR). Some of the proteins also stabilize  
 CC duplexes during polymerase reactions or improve exonuclease reactions,  
 CC for example RFA also improves specificity of nucleic acid/protein  
 CC interaction and PCNA improves polymerase-mediated repair processes and  
 CC hybridization reactions. Nucleic acids encoding the archaeal polypeptides  
 CC are used for recombinant production of proteins, and fragments of the  
 CC nucleic acid as probes and primers for screening related sequences. The  
 CC accessory proteins increase accuracy and efficiency of polymerase  
 CC reactions, allow use of lower denaturation and extension temperatures  
 CC (possibly isothermal processing), and improve synthesis of long targets.  
 CC The present sequence represents a P. furiosus recombinant helicase 7  
 XX  
 SQ Sequence 1278 AA;

Query Match 84.1%; Score 37; DB 4; Length 1278;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLEFDA 8  
 DB 650 YLQLEFDA 657  
 :|||||

RESULT 5  
 ADB06652  
 ID ADB06652 standard; protein; 282 AA.  
 XX  
 AC ADB06652;  
 XX

DT 20-NOV-2003 (first entry)  
 XX  
 DE Alloicoccus otitis antigenic protein SEQ ID NO:592.  
 XX  
 KW Alloicoccus otitidis; antigenic protein; immunogenic; immunisation;  
 KW gene therapy; Gram-positive bacterium; infection.  
 XX  
 OS Alloicoccus otitis.  
 XX  
 PN WO2003048304-A2.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PF 25-NOV-2002; 2002WO-US036123.  
 XX  
 PR 29-NOV-2001; 2001US-0333777P.  
 PR 18-NOV-2002; 2002US-0426742P.  
 XX  
 PA (AMHP ) WYETH HOLDINGS CORP.  
 XX  
 PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;  
 XX  
 DR WPI; 2003-505284/47.  
 DR N-PSDB; ADB06651.  
 XX  
 PT New Alloicoccus otitidis polynucleotides and polypeptides, useful for  
 PT treating and diagnosing diseases, drug screening assays and monitoring of  
 PT effects during drug clinical trials.  
 XX  
 PS Claim 33; SEQ ID NO 592; 1019pp; English.  
 XX  
 CC The present invention describes an isolated polynucleotide (I) of  
 CC Alloicoccus otitidis genomic DNA, which encodes an antigenic protein.  
 CC Alloicoccus otitidis is a Gram-positive bacterium. Also described: (1)  
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an  
 CC expression vector comprising the novel isolated polynucleotide (I), its  
 CC complement, degenerate variant or fragment; (3) a genetically engineered  
 CC host cell, transfected, transformed or infected with the vector of (2);

CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
 CC composition comprising the polypeptide, its complement, biological  
 CC equivalent or fragment, or the polynucleotide that is comprised in the  
 CC expression vector; (6) a pharmaceutical composition comprising the  
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)  
 CC immunising against *Alloiooccus* otitidis by administering to a host the  
 CC immunogenic composition; (9) detecting and/or identifying *Alloiooccus*  
 CC otitidis in the biological sample; (10) a kit comprising a container  
 CC containing the novel polynucleotide, its degenerate variant or fragment,  
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the  
 CC genetically engineered host cell under conditions suitable to produce the  
 CC polypeptide from the culture. (1) can be used in gene therapy. The  
 CC polynucleotides, polypeptides, antibodies and compositions of the present  
 CC invention can be used for treating and diagnosing diseases, drug  
 CC screening assays and monitoring of effects during drug clinical trials.  
 CC The polynucleotides are useful for expressing and detecting *Alloiooccus*  
 CC otitidis. The present sequence represents an *Alloiooccus* otitidis  
 CC antigen protein from the present invention.

XX SQ Sequence 282 AA;

Query Match 79.5%; Score 35; DB 6; Length 282;  
 Best Local Similarity 77.8%; Pred. No. 72;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFDDAV 9  
 ||||| |:  
 Db 184 FLQLEIDSV 192

# RESULT 6

ADB06654  
 ID ADB06654 standard; protein; 285 AA.

AC ADB06654;

XX 20-NOV-2003 (first entry)

DT *Alloiooccus* otitis antigenic protein SEQ ID NO:594.

DE *Alloiooccus* otitidis; antigenic protein; immunogenic; immunisation;  
 KW gene therapy; Gram-positive bacterium; infection.

XX *Alloiooccus* otitis.

OS WO2003048304-A2.

FN 12-JUN-2003.

PD 25-NOV-2002; 2002WO-US036123.

PF 29-NOV-2001; 2001US-0333777P.

XX 18-NOV-2002; 2002US-0426742P.

PR (AMHP ) WYETH HOLDINGS CORP.

XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;

PI WPI; 2003-505284/47.

XX N-PSDB; ADB06653.

XX New *Alloiooccus* otitidis polynucleotides and polypeptides, useful for  
 PT treating and diagnosing diseases, drug screening assays and monitoring of  
 FT effects during drug clinical trials.

XX Claim 33; SEQ ID NO 594; 1019pp; English.

XX The present invention describes an isolated polynucleotide (I) of  
 CC *Alloiooccus* otitidis genomic DNA, which encodes an antigenic protein.  
 CC *Alloiooccus* otitidis is a Gram-positive bacterium. Also described: (1)  
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an  
 CC expression vector comprising the novel isolated polynucleotide (I), its

CC complement, degenerate variant or fragment; (3) a genetically engineered  
 CC host cell, transfected, transformed or infected with the vector of (2);  
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
 CC composition comprising the polypeptide, its complement, biological  
 CC equivalent or fragment, or the polynucleotide that is comprised in the  
 CC expression vector; (6) a pharmaceutical composition comprising the  
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)  
 CC immunising against *Alloiooccus* otitidis by administering to a host the  
 CC immunogenic composition; (9) detecting and/or identifying *Alloiooccus*  
 CC otitidis in the biological sample; (10) a kit comprising a container  
 CC containing the novel polynucleotide, its degenerate variant or fragment,  
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the  
 CC genetically engineered host cell under conditions suitable to produce the  
 CC polypeptide from the culture. (1) can be used in gene therapy. The  
 CC polynucleotides, polypeptides, antibodies and compositions of the present  
 CC invention can be used for treating and diagnosing diseases, drug  
 CC screening assays and monitoring of effects during drug clinical trials.  
 CC The polynucleotides are useful for expressing and detecting *Alloiooccus*  
 CC otitidis. The present sequence represents an *Alloiooccus* otitidis  
 CC antigen protein from the present invention.

XX SQ Sequence 285 AA;

Query Match 79.5%; Score 35; DB 6; Length 285;  
 Best Local Similarity 77.8%; Pred. No. 73;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFDDAV 9  
 ||||| |:  
 Db 187 FLQLEIDSV 195

# RESULT 7

ABB59769

ID ABB59769 standard; protein; 1612 AA.

XX ABB59769;

AC 26-MAR-2002 (first entry)

DT *Drosophila melanogaster* polypeptide SEQ ID NO 6099.

DE *Drosophila*; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX *Drosophila melanogaster*.

OS WO200171042-A2.

FN 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

XX N-PSDB; ABL03872.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
 FT interactions.

XX Disclosure; SEQ ID NO 6099; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 1612 AA;

Query Match 79.5%; Score 35; DB 4; Length 1612;  
Best Local Similarity 66.7%; Pred. No. 4.6e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLQEFDAV 9  
Db 254 YLELEFDAL 262  
:|||||:

RESULT 8  
ID ABM74402  
XX ABM74402 standard; protein; 217 AA.  
AC ABM74402;  
XX  
DT 17-OCT-2003 (first entry)  
XX  
DE DNA clone originating in barley containing SNP sequence #812.  
XX  
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  
XX  
OS Hordeum vulgare.  
XX  
PN WO2003057877-A1.  
XX  
PD 17-JUL-2003.  
XX  
PF 16-DEC-2002; 2002WO-IB005403.  
XX  
PR 20-DEC-2001; 2001JP-00387059.  
PR 20-DEC-2001; 2001JP-00387131.  
PR 20-DEC-2001; 2001JP-00403299.  
PR 20-DEC-2001; 2001JP-00403300.  
PR 27-SEP-2002; 2002JP-00327515.  
XX  
XX (UYN1-) UNIV JAPAN OKAYAMA.  
XX  
PI Sato K, Takeda K, Kohara Y;  
XX  
XX WPI; 2003-587127/55.  
XX  
PT Single nucleotide polymorphism sites in barley varieties and DNA  
PT sequences containing them for analysis and identification of barley  
PT varieties and production of barley transformants with desired  
PT characteristics.  
XX  
PS Disclosure; SEQ ID XX; 284pp; Japanese.  
XX  
CC The present invention relates to oligonucleotide clones originating in  
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
CC varieties, identification of particular varieties and genotype-phenotype  
CC analysis, isolation of specific genes and creation of new varieties by  
CC transformation of barley varieties with them and production of new barley  
CC varieties with desired properties. The present sequence represents an  
CC oligonucleotide clone sequence featured in the specification. The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published-pct-sequences  
XX  
SQ Sequence 217 AA;

Query Match 75.0%; Score 33; DB 7; Length 217;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLQEFDAV 9  
Db 44 FLQIKFNAV 52  
|||||:

RESULT 9  
ADP98863  
ID ADP98863 standard; protein; 374 AA.  
XX  
AC ADP98863;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE C. albicans specific gene, orf6.4105, protein sequence.  
XX  
KW Diploid fungal cell; allele; gene disruption cassette;  
KW promoter replacement fragment; antifungal; fungicide; gene therapy;  
KW infection; Candida albicans.  
XX  
OS Candida albicans.  
XX  
PN WO2004056965-A2.  
XX  
PD 08-JUL-2004.  
XX  
PF 19-DEC-2003; 2003WO-US040618.  
XX  
PR 19-DEC-2002; 2002US-0434832P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX (ELIT-) ELITRA CANADA LTD.  
XX  
PI Roemer T, Jiang B, Boone C, Bussey H;  
XX  
XX WPI; 2004-500296/47.  
DR N-PSDB; ADP98853.  
XX  
XX Constructing a strain of diploid fungal cells in which both alleles of a  
PT gene are modified comprises modifying the alleles of a gene in the fungal  
PT cells by recombination using a gene disruption cassette and a promoter  
PT replacement fragment.  
XX  
PS Claim 44; SEQ ID NO 7038; 163pp; English.  
XX  
CC The invention relates to a novel method for constructing a strain of  
CC diploid fungal cells in which both alleles of a gene are modified. The  
CC method comprises modifying the alleles of a gene in diploid fungal cells  
CC by recombination using a gene disruption cassette and a promoter  
CC replacement fragment. The invention further comprises: assembling a  
CC collection of diploid fungal cells each of which comprises modified  
CC alleles of a different gene; a strain of diploid fungal cells comprising  
CC modified alleles of a gene, where the first allele of the gene is  
CC inactivated by a gene disruption cassette comprising a nucleotide  
CC sequence encoding an expressible selectable marker; and the expression of  
CC the second allele of the gene is regulated by a heterologous promoter  
CC that is operably linked to the coding region of the second allele of the  
CC gene, and where the gene encodes the polypeptide mentioned above; a  
CC collection of diploid fungal strains comprising the diploid strains cited  
CC above, where substantially all the different genes that encode the above  
CC amino acid sequences are modified and are present in different diploid  
CC strains in the collection; a nucleic acid molecule microarray comprising  
CC nucleic acid molecules, where each nucleic acid molecule comprises a  
CC nucleotide sequence that is hybridizable to a target nucleotide sequence  
CC comprising any of the 310 nucleotide sequences listed in the  
CC specification (ADP98516-ADP98825); identifying a gene that is essential  
CC to the survival or growth of a fungus, that contributes to the virulence  
CC and/or pathogenicity of a fungus, or that contributes to the resistance  
CC of a diploid fungus to an antifungal agent; identifying an antifungal  
CC agent that inhibits the growth of a diploid fungus, or a therapeutic  
CC agent for treatment of a mammalian disease; correlating changes in the

CC levels of proteins or gene transcripts with the inhibition of growth or  
 CC proliferation of a diploid fungal cell; a purified or isolated nucleic  
 CC acid molecule comprising a nucleotide sequence encoding a gene product  
 CC required for proliferation of Candida albicans, where the gene product  
 CC consists of any of the above-mentioned amino acid sequences; a vector  
 CC comprising a promoter operably linked to the nucleic acid molecule cited  
 CC above; a host cell containing the vector; a purified or isolated  
 CC polypeptide comprising any of the 61 amino acid sequences given in the  
 CC specification (ADP96718-ADP96778); a fusion protein comprising a fragment  
 CC of a first polypeptide fused to a second polypeptide, the fragment  
 CC consisting of at least 6 consecutive residues of any of ADP98826-ADP99135  
 CC ; producing a polypeptide; identifying a compound which modulates the  
 CC activity of a gene product encoded by a nucleic acid comprising any of  
 CC ADP98516-ADP98825; eliciting an immune response in an animal; a strain of  
 CC Candida albicans, where a first allele of a gene comprising any of  
 CC ADP98516-ADP98825 is inactive and a second allele of the gene is under  
 CC the control of a heterologous promoter; identifying a compound or binding  
 CC partner that binds to the polypeptide comprising any of ADP98826-  
 CC ADP99135, or its fragment; identifying a compound having the ability to  
 CC inhibit growth or proliferation of Candida albicans; inhibiting growth or  
 CC proliferation of Candida albicans cells; manufacturing an antimycotic  
 CC compound; treating an infection of a subject by Candida albicans;  
 CC preventing or containing contamination of an object by Candida albicans,  
 CC or for preventing or inhibiting formation on a surface of a biofilm  
 CC comprising Candida albicans; a pharmaceutical composition comprising a  
 CC therapeutic amount of an agent which reduces the activity or level of a  
 CC gene product encoded by a nucleic acid comprising any of ADP98516-  
 CC ADP98825 in a pharmaceutical carrier; an antibody preparation which binds  
 CC the polypeptide; methods for evaluating a compound against a target gene  
 CC product encoded by any of ADP98516-ADP98825; identifying an antimycotic  
 CC compound; a computer or a computer readable medium that comprises at  
 CC least one of the nucleotide sequences mentioned in the specification or  
 CC at least one amino acid sequence selected from ADP98826-ADP99135; a  
 CC method assisted by a computer for identifying a putatively essential gene  
 CC of a fungus; and a protein array comprising proteins, where at least one  
 CC sequence selected from ADP98516-ADP98825. The novel methods and  
 CC compositions have fungicide activity. The compositions may be used in  
 CC gene therapy. The composition and methods are useful for drug screening  
 CC with Candida albicans. These may also be used for treating infections associated  
 CC useful for identification and validation of gene products as effective  
 CC targets for therapeutic intervention, for identifying and validating gene  
 CC products as effective targets for therapeutic intervention, and for  
 CC collecting identified essential genes. This sequence represents the  
 CC protein of a Candida albicans fungal specific gene of the invention.  
 CC NOTE: This sequence was downloaded from an electronic sequence listing  
 CC provided on the WIPO website.

SQ Sequence 374 AA;

Query Match 75.0%; Score 33; DB 8; Length 374;  
 Best Local Similarity 85.7%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLEFDDAV 9

Db 293 QLEFDIAI 299

RESULT 10  
 ABP06827

ID ABP06827 standard; protein; 72 AA.

XX AC ABP06827;

DT 24-JUN-2002 (first entry)

XX Human ORFX protein sequence SEQ ID NO:13636.

DE Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;

KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis.

OS Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US010836.

XX 30-MAY-2000; 2000US-0206132P.

XX 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX WPI; 2002-106308/14.

XX N-PSDB; ABN22579.

XX Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure; SEQ ID NO 13636; 1037pp; English.

XX The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 72 AA;

Query Match 72.7%; Score 32; DB 5; Length 72;

Best Local Similarity 66.7%; Pred. No. 69;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLFDDAV 9

Db 54 FLQLDWDVV 62

RESULT 11

ABU70578

ID ABU70578 standard; protein; 173 AA.

XX AC ABU70578;

DT 10-JUN-2003 (first entry)

XX DE Human adipocyte Selected Interacting domain, SID, #205.  
XX DE Human; prey; adipocyte; SID: selected interacting domain; anorectic;  
XX KW antidiabetic; protein-protein interaction; diabetes;  
XX KW yeast 2-hybrid assay; metabolic disorder; obesity.  
XX OS Homo sapiens.  
XX PN WO200286122-A2.  
XX PD 31-OCT-2002.  
XX PF 14-MAR-2002; 2002WO-EP003768.  
XX PR 14-MAR-2001; 2001US-0275734P.  
XX PA (HYBR-) HYBRIGENICS.  
XX PI Legrain P, Daviet L;  
XX PI WPI; 2003-103412/09.  
XX DR N-PSDB; ACA57122.  
XX XX New complex between two interacting proteins in adipocyte cells, useful  
PT for identifying selected interacting domains that modulate protein  
PT interactions, or for preventing or treating metabolic disorders such as  
PT obesity or diabetes.  
XX PS Claim 6; Page 175; 382pp; English.  
XX XX The invention relates to a complex between two interacting proteins in  
CC adipocyte cells, given in the specification. The proteins are identified  
CC by selecting a bait protein from a known adipocyte marker and then  
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by  
CC members of an adipocyte cDNA library. The proteins are designated SID  
CC (RTM) (selected interacting domains) proteins. Also included are a  
CC recombinant host cell expressing a polypeptide in the adipocyte cells, a  
CC polypeptide of the complex, selecting a modulating compound in adipocyte  
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid  
CC sequences given in the specification (including its fragment or variant),  
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences  
CC given in the specification (including its fragment or variant), a vector  
CC comprising the SID (RTM) polynucleotide, a recombinant host cell  
CC comprising the vector, a protein chip comprising the polypeptides and a  
CC record comprising all or part of the data, listed in the specification.  
CC The complex, polypeptides, polynucleotides and compounds are useful for  
CC preventing or treating metabolic disorders such as obesity or diabetes.  
CC The polynucleotides are useful as probes or primers. The complex is  
CC particularly useful for identifying selected interacting domains (SID  
CC (RTM)) for screening drugs that modulate the protein interaction, thus  
CC exhibiting the therapeutic effect. The present sequence represents a SID  
CC (prey) protein of the invention  
XX SQ Sequence 173 AA;  
Query Match 72.7%; Score 32; DB 6; Length 173;  
Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 FQLEFD 7  
Db 141 FQLKFD 147  
RESULT 12  
ABU21843  
ID ABU21843 standard; protein; 225 AA.  
XX AC ABU21843;  
XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by prokaryotic essential gene #7370.  
XX DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX KW Burkholderia fungorum.  
XX OS WO200277183-A2.  
XX PN 03-OCT-2002.  
XX PD 21-MAR-2002; 2002WO-US009107.  
XX PF 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
XX DR N-PSDB; ACA25713.  
XX XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 49767; 1766pp; English.  
XX XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than S. aureus, S. typhimurium,  
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 225 AA;  
Query Match 72.7%; Score 32; DB 6; Length 225;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 FQLEFD 9  
Db 141 FQLKFD 147







RESULT 15  
ABO78318  
ID ABO78318 standard; protein; 419 AA.  
XX  
AC ABO78318;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #10493.  
XX  
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
DR WPI; 2003-615309/58.  
DR N-PSDB; ABD11889.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 27064; 455pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 419 AA;  
  
Query Match 72.7%; Score 32; DB 7; Length 419;  
Best Local Similarity 75.0%; Pred. No. 4.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 LQLEFDV 9  
Db 226 LQVEFDAL 233  
  
Search completed: December 30, 2004, 20:42:58  
Job time : 12.4962 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:30:54 ; Search time 1.97229 Seconds  
(without alignments)  
439.058 Million cell updates/sec

Title: US-10-017-327-5  
Perfect score: 44  
Sequence: 1 FLQLEFDVAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	79.5	129	G82526	ATP-dependent DNA
2	33	75.0	129	G64232	hypothetical prote
3	33	75.0	430	B82096	conserved hypothet
4	32	72.7	154	G95328	hypothetical prote
5	32	72.7	228	T12000	cytochrome-c oxida
6	32	72.7	228	T09802	cytochrome-c oxida
7	32	72.7	229	S50328	probable metal-dep
8	32	72.7	274	G97034	prostaglandin-endo
9	32	72.7	599	JH0259	prostaglandin-endo
10	32	72.7	599	A29947	prostaglandin-endo
11	32	72.7	600	S00561	prostaglandin-endo
12	32	72.7	600	A28960	prostaglandin-endo
13	32	72.7	602	S39782	cyclooxygenase 1 -
14	32	72.7	602	A35564	prostaglandin-endo
15	32	72.7	602	S89198	prostaglandin G/H
16	32	72.7	756	F83704	homocysteine methyl
17	31	70.5	204	A32252	probable phosphoes
18	31	70.5	210	A40979	temperature shock
19	31	70.5	321	G70415	nucleotide sugar e
20	31	70.5	360	T06786	6a-hydroxymaackiai
21	31	70.5	384	A86324	protein F14D16.19
22	31	70.5	390	F81393	probable altronate
23	31	70.5	400	B90139	phosphomethylpyrim
24	31	70.5	578	T44444	hypothetical prote
25	31	70.5	875	T19678	hypothetical prote
26	31	70.5	1042	G64514	type I restriction
27	31	70.5	1852	JC5546	chitin synthase (E
28	30	68.2	100	A20773	hypothetical prote
29	30	68.2	155	E84198	hypothetical prote

30	30	68.2	211	2	D69888	micrococcal nuclea
31	30	68.2	226	2	G45170	cytochrome-c oxida
32	30	68.2	229	2	G90730	antitermination pr
33	30	68.2	237	2	B81870	probable membrane
34	30	68.2	257	2	D85581	hypothetical prote
35	30	68.2	272	1	G82076	3',5'-cyclic-nucle
36	30	68.2	280	2	T38816	hypothetical prote
37	30	68.2	290	2	B97100	pyridoxal kinase r
38	30	68.2	296	2	T12770	probable endonucle
39	30	68.2	299	2	T23932	hypothetical prote
40	30	68.2	300	2	C89857	formyltetrahydrofo
41	30	68.2	301	2	C85574	probable lysR-like
42	30	68.2	301	2	C90723	probable transcrip
43	30	68.2	365	2	T04247	hypothetical prote
44	30	68.2	372	2	T01551	receptor kinase ho
45	30	68.2	420	2	T51874	hypothetical prote

ALIGNMENTS

RESULT 1

G82526  
ATP-dependent DNA helicase XF2680 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: G82526  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: G82526  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-880 <SIM>  
A:Cross-references: UNIPROT:Q9PA40; GB:AE004074; GB:AE003849; NID:g9107918; PIDN:AAF8547  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Carrer, H.  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.  
as-Neto, E.; Docena, C.; El-Dorzy, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2680  
C:Superfamily: helicase II

Query Match 79.5%; Score 35; DB 2; Length 680;  
Best Local Similarity 77.8%; Pred. No. 22;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLEFDVAV 9  
|||:|  
Db 114 FLQIEHDAV 122

RESULT 2

G64232  
hypothetical protein MG296 - Mycoplasma genitalium  
C:Species: Mycoplasma genitalium  
C>Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: G64232  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudak, D.M.; Phillips, C.A.; Merrick, J.

C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A>Title: The minimal gene complement of Mycoplasma genitalium.  
 A;Reference number: A64200; MUID:96026346; PMID:7569993  
 A;Accession: G64232  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-129 <TIGR>  
 A;Cross-references: UNIPROT:P47538; GB:U39710; GB:L43967; NID:g1045989; PID:g1045994; TIGR:P47538  
 A;Experimental source: strain G-37  
 C;Genetics:  
 A;Genetic code: SGC3

Query Match 75.0%; Score 33; DB 2; Length 129;  
 Best Local Similarity 77.8%; Pred. No. 9;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLQLEFDV 9  
 |||||  
 Db 11 FLQTEFQV 19

RESULT 3  
 B82096  
 conserved hypothetical protein VC2278 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C;Species: Vibrio cholerae  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C;Accession: B82096  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Charles, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.F.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301  
 A;Accession: B82096  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-430 <HBI>  
 A;Cross-references: UNIPROT:Q9KPT4; GB:AE004299; GB:AE003852; NID:g9656835; PIDN:AAF9542  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC2278  
 A;Map position: 1  
 C;Superfamily: conserved hypothetical protein HI0125

Query Match 75.0%; Score 33; DB 2; Length 430;  
 Best Local Similarity 66.7%; Pred. No. 35;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLEFDV 9  
 |||||  
 Db 227 FMQLDFSAV 235

RESULT 4  
 G95328  
 hypothetical protein SMA0994 [imported] - Sinorhizobium meliloti (strain 1021) magaplasma  
 C;Species: Sinorhizobium meliloti  
 C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
 C;Accession: G95328  
 R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows, J.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
 A;Reference number: A95262; MUID:21396509; PMID:11481432  
 A;Accession: G95328  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-154 <KUR>  
 A;Cross-references: UNIPROT:Q922F1; GB:AE006469; PIDN:AAK65193.1; PID:g14523639; GSPDB:G  
 A;Experimental source: strain 1021, megaplasmaid pSymA  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Bows, J.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Puhler, A.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.  
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A;Reference number: A96039; MUID:21368234; PMID:11474104  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: SMA0994  
 A;Genome: plasmid

Query Match 72.7%; Score 32; DB 2; Length 154;  
 Best Local Similarity 87.5%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQLEFDV 9  
 |||||  
 Db 109 LQLAFDAV 116

RESULT 5  
 T12000  
 cytochrome-c oxidase (EC 1.9.3.1) chain II - Anopheles quadrimaculatus A mitochondrion  
 C;Species: mitochondrion Anopheles quadrimaculatus A  
 C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 18-Aug-2003  
 C;Accession: T12000  
 R;Cockburn, A.F.; Mitchell, S.E.; Seawright, J.A.  
 Arch. Insect Biochem. Physiol. 14, 31-36, 1990  
 A>Title: Cloning of the mitochondrial genome of Anopheles quadrimaculatus.  
 A;Reference number: Z17375; MUID:92190510; PMID:2134168  
 A;Accession: T12000  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-228 <COC>  
 A;Cross-references: EMBL:L04272; NID:g342501; PID:g507281; PIDN:AAA93542.1  
 A;Experimental source: strain Orlando  
 C;Genetics:  
 A;Gene: COXII  
 A;Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase chain II  
 C;Keywords: copper; electron transfer; membrane-associated complex; mitochondrial inner membrane; copper 1 (His, Cys, Cys, Met) #status predicted  
 F;161,196,200,207/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
 F;196,198,200,204/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
 F;198/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 72.7%; Score 32; DB 2; Length 228;  
 Best Local Similarity 75.0%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQLEFDA 8  
 |||||  
 Db 113 FLNLEFDS 120

RESULT 6  
 T09802  
 cytochrome-c oxidase (EC 1.9.3.1) chain II - African malaria mosquito mitochondrion  
 C;Species: mitochondrion Anopheles gambiae (African malaria mosquito)  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C;Accession: T09802  
 R;Beard, C.B.; Hamm, D.M.; Collins, F.H.  
 Insect Mol. Biol. 2, 103-124, 1993  
 A>Title: The mitochondrial genome of the mosquito Anopheles gambiae: DNA sequence, genome organization, and predicted functions  
 A;Reference number: Z16863; MUID:97242550; PMID:9087549  
 A;Accession: T09802  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-228 <BEA>  
 A;Cross-references: UNIPROT:P34840; EMBL:L20934; NID:g309056; PIDN:AA12192.1; PID:g309056  
 A;Experimental source: strain G3  
 C;Genetics:  
 A;Gene: mitochondrion  
 A;Genetic code: SGC4

C;Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase d  
F;9-214/Domain: cytochrome-c oxidase chain II homology <CO2>  
F;161,196,200,207/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
F;196,198,200,204/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
F;198/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 72.7%; Score 32; DB 2; Length 228;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFDA 8  
|||:|:|:  
DB 113 FLNLEFDS 120

RESULT 7  
S50328  
cytochrome-c oxidase (EC 1.9.3.1) chain II - Katharina tunicata mitochondrion  
C;Species: mitochondrion Katharina tunicata  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S50328  
R;Boore, J.L.; Brown, W.M.  
Genetics 138, 423-443, 1994  
A;Title: Complete DNA sequence of the mitochondrial genome of the black chiton, Katharin  
A;Reference number: S50327; MUID:95129806; PMID:7828825  
A;Accession: S50328  
A;Molecule type: DNA  
A;Residues: 1-229 <BO>  
A;Cross-references: UNIPROT:Q37534; EMBL:U09810; NID:9557273; PIDN:AAC48365.1; PID:95572  
C;Genetics:  
A;Genome: mitochondrion  
A;Genetic code: SGC4  
C;Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase d  
C;Keywords: copper; electron transfer; heme; membrane-associated complex; mitochondrial  
ein

F;9-214/Domain: cytochrome-c oxidase chain II homology <CO2>  
F;161,196,200,207/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
F;196,198,200,204/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
F;198/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 72.7%; Score 32; DB 2; Length 229;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFDA 8  
|||:|:|:  
DB 113 FLNLEFDS 120

RESULT 8  
G97034  
probable metal-dependent phosphoesterase (PHP family), YciV ortholog [imported] - Clostr  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: G97034  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: G97034  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-274 <KUR>  
A;Cross-references: UNIPROT:Q97K29; GB:AE001437; PIDN:AAK99066.1; PID:gl5024008; GSPDB:G  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC1092  
C;Superfamily: hypothetical protein H11400

Query Match 72.7%; Score 32; DB 2; Length 274;  
Best Local Similarity 55.6%; Pred. No. 34;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEFDAV 9  
|||:|:|:  
DB 205 FLKLDFFDGI 213

RESULT 9  
JH0259  
prostaglandin-endoperoxide synthase (EC 1.14.99.1) 1 precursor - human  
N;Alternate names: cyclooxygenase; prostaglandin G/H synthase; prostaglandin H synthase  
N;Contains: prostaglandin-endoperoxide synthase 1, splice form 2  
C;Species: Homo sapiens (man)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C;Accession: JH0259; PH0225; A39937; B38146; A38146; S50181; A36746; S69169  
R;Takahashi, Y.; Ueda, N.; Yoshimoto, T.; Yamamoto, S.; Yokoyama, C.; Miyata, A.; Tanabe  
Biochem. Biophys. Res. Commun. 182, 433-438, 1992  
A;Title: Immunofluorescence purification and cDNA cloning of human platelet prostaglandin en  
A;Reference number: JH0259; MUID:92134251; PMID:1734857  
A;Accession: JH0259  
A;Molecule type: mRNA  
A;Residues: 1-599 <TAK1>  
A;Cross-references: UNIPROT:P23219; GB:S78220; NID:g243971; PIDN:AAB1215.1; PID:g243972  
A;Accession: PH0225  
A;Molecule type: protein  
A;Residues: 24-31 <TAK2>  
R;Funk, C.D.; Funk, L.B.; Kennedy, M.E.; Pong, A.S.; Fitzgerald, G.A.  
FASEB J. 5, 2304-2312, 1991  
A;Title: Human platelet/erythrocyte cell prostaglandin G/H synthase: cDNA cloning, e  
A;Reference number: A39937; MUID:91317397; PMID:1907252  
A;Accession: A39937  
A;Molecule type: mRNA  
A;Residues: 1-599 <FUN>  
A;Cross-references: GB:M59979; NID:g189886; PIDN:AAA03630.1; PID:g189887  
R;Diaz, A.; Reginato, A.M.; Jimenez, S.A.  
J. Biol. Chem. 267, 10816-10822, 1992  
A;Title: Alternative splicing of human prostaglandin G/H synthase mRNA and evidence of d  
nd tumor necrosis factor alpha.  
A;Reference number: A38146; MUID:92268138; PMID:1587858  
A;Accession: B38146  
A;Molecule type: mRNA  
A;Residues: 1-599 <DIA1>  
A;Cross-references: GB:S36271; NID:g249625; PIDN:AAB22217.1; PID:g249626  
A;Experimental source: lung fibroblast  
A;Note: sequence extracted from NCBI backbone (NCBIN:103945, NCBIP:103946)  
A;Accession: A38146  
A;Molecule type: mRNA  
A;Residues: 1-395,433-599 <DIA2>  
A;Cross-references: GB:S36219; NID:g249623; PIDN:AAB22216.1; PID:g249624  
A;Experimental source: lung fibroblast  
A;Note: sequence extracted from NCBI backbone (NCBIN:103825, NCBIP:103826)  
R;Barnett, J.; Chow, J.; Ives, D.; Chiou, M.; Mackenzie, R.; Osen, E.; Nguyen, B.; Tsing  
Biochim. Biophys. Acta 1209, 130-139, 1994  
A;Title: Purification, characterization and selective inhibition of human prostaglandin  
A;Reference number: S50181; MUID:95035046; PMID:7947975  
A;Accession: S50181  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 24-32 <BAK>  
R;Yokoyama, C.; Tanabe, T.  
Biochem. Biophys. Res. Commun. 165, 888-894, 1989  
A;Title: Cloning of human gene encoding prostaglandin endoperoxide synthase and primary  
A;Reference number: A36746; MUID:90088508; PMID:2512924  
A;Accession: A36746  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-11,'L',13-112,'L',114-377,'T',379-599 <YOK>  
A;Cross-references: GB:M31822; NID:g189898; PIDN:AAA36439.1; PID:g387018  
R;Ren, Y.; Loose-Mitchell, D.S.; Kulmacz, R.J.  
Arch. Biochem. Biophys. 316, 751-757, 1995  
A;Title: Prostaglandin H synthase-1: evaluation of C-terminus function.  
A;Reference number: S69169; MUID:95168861; PMID:7864630  
A;Accession: S69169

A:Molecule type: protein  
A:Residues: 585-599 <REN>  
C:Genetics:  
A:Gene: GDB:PTGS1  
A:Cross-references: GDB:128070; OMIM:176805  
A:Map position: 9q32-q33.3  
C:Function:  
A:Description: catalyzes the oxidative cyclization by oxygen of arachidonic acid to prostaglandin biosynthesis  
A:Pathway: prostaglandin biosynthesis  
C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
C:Keywords: alternative splicing; chromoprotein; endoplasmic reticulum; glycoprotein; he  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-599/Product: prostaglandin-endoperoxide synthase 1 #status experimental <MARI>  
F:24-395,433-599/Product: prostaglandin-endoperoxide synthase 1, splice form 2 #status p  
F:35-68/Domain: EGF homology <EGF>  
F:103,143,409/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:308/Binding site: heme iron (His) (axial ligand) #status predicted  
F:384,529/Active site: Tyr, Ser #status predicted

Query Match 72.7%; Score 32; DB 2; Length 599;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFD 7  
|||:|  
Db 355 FLQLKFD 361

RESULT 10  
A29947  
prostaglandin-endoperoxide synthase (EC 1.14.99.1) precursor - sheep  
N:Alternate names: prostaglandin endoperoxide synthetase  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 09-Jul-2004  
C:Accession: A29947  
R:Merlie, J.P.; Fagan, D.; Mudd, J.; Needleman, P.  
J. Biol. Chem. 263, 3550-3553, 1988  
A:Title: Isolation and characterization of the complementary DNA for sheep seminal vesic  
A:Reference number: A29947; MUID:88153641; PMID:2831188  
A:Accession: A29947  
A:Molecule type: mRNA  
A:Residues: 1-599 <MER>  
A:Cross-references: UNIPROT:P05979; GB:M18243; NID:g165843; PIDN:AAA31511.1; PID:g165844  
C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
C:Keywords: oxidoreductase  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-599/Product: prostaglandin-endoperoxide synthase #status predicted <MAT>  
F:35-68/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 599;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFD 7  
|||:|  
Db 355 FLQLKFD 361

RESULT 11  
S00561  
prostaglandin-endoperoxide synthase (EC 1.14.99.1) precursor - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 10-Dec-1999  
C:Accession: S00561  
R:Yokoyama, C.; Takai, T.; Tanabe, T.  
FEBS Lett. 231, 347-351, 1988  
A:Title: Primary structure of sheep prostaglandin endoperoxide synthase deduced from cDN  
A:Reference number: S00561; MUID:88196421; PMID:3129310  
A:Accession: S00561  
A:Molecule type: mRNA  
A:Residues: 1-600 <YOK>  
A:Cross-references: EMBL:Y00750; NID:g1361; PIDN:CAA68719.1; PID:g1362  
A:Note: part of this sequence, including the amino end of the mature protein, was confir

A:Note: 97-His, 164-Gly, 456-Gln, 520-Gln, 520-Lys, and 525-Ile were also found  
C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
C:Keywords: oxidoreductase  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-600/Product: prostaglandin-endoperoxide synthase #status experimental <MAT>  
F:36-69/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 600;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFD 7  
|||:|  
Db 356 FLQLKFD 362

RESULT 12  
A28960

prostaglandin-endoperoxide synthase (EC 1.14.99.1) G/H precursor - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C:Accession: A28960  
R:DeWitt, D.L.; Smith, W.L.  
Proc. Natl. Acad. Sci. U.S.A. 85, 1412-1416, 1988  
A:Title: Primary structure of prostaglandin G/H synthase from sheep vesicular gland dete  
A:Reference number: A28960; MUID:88144447; PMID:3125548  
A:Accession: A28960  
A:Molecule type: mRNA  
A:Residues: 1-600 <DEW>  
A:Cross-references: UNIPROT:P05979; GB:J03599; NID:g166035; PIDN:AAA31576.1; PID:g166036  
C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
C:Keywords: oxidoreductase  
F:36-69/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 600;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFD 7  
|||:|  
Db 356 FLQLKFD 362

RESULT 13  
S39782

cyclooxygenase 1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 10-Dec-1999  
C:Accession: S39782  
R:Feng, L.; Sun, W.; Xia, Y.; Tang, W.W.; Channugam, P.; Soyoola, E.; Wilson, C.B.; Hwang  
Arch. Biochem. Biophys. 307, 361-368, 1993  
A:Title: Cloning two isoforms of rat cyclooxygenase: differential regulation of their exp  
A:Reference number: S39782; MUID:94099619; PMID:8274023  
A:Accession: S39782  
A:Molecule type: mRNA  
A:Residues: 1-602 <FEN>  
A:Cross-references: GB:S67721; NID:g460555; PIDN:AAB29400.1; PID:g460556  
C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
F:38-71/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 602;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFD 7  
|||:|  
Db 358 FLQLKFD 364

RESULT 14  
A35564

prostaglandin-endoperoxide synthase (EC 1.14.99.1) precursor - mouse  
C:Species: Mus musculus (house mouse)

C;Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
 C;Accession: A35564  
 R;Dewitt, D.L.; El-Harith, E.A.; Kraemer, S.A.; Andrews, M.J.; Yao, E.F.; Armstrong, R.I.  
 J. Biol. Chem. 265, 5192-5198, 1990  
 A;Title: The aspirin and heme-binding sites of ovine and murine prostaglandin endoperoxide synthase  
 A;Reference number: A35564; MUID:90203007; PMID:2108169  
 A;Accession: A35564  
 A;Molecule type: mRNA

A;Residues: 1-602 <DEW>  
 A;Cross-references: UNIPROT:P22437; GB:M34141; NID:G200302; PIDN:AAA39913.1; PID:G200302  
 C;Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 C;Keywords: oxidoreductase  
 F;1-26/Domain: signal sequence #status predicted <SIG>  
 F;27-602/Product: prostaglandin-endoperoxide synthase #status predicted <MAT>  
 F;38-71/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLEFD 7  
 |||||  
 Db 358 FLQLKFD 364

## RESULT 15

S69198  
 prostaglandin G/H synthase 1 - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C;Accession: S69198; S69199

R;Kitzler, J.W.  
 submitted to the EMBL Data Library, December 1994  
 A;Reference number: S69198  
 A;Accession: S69198  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-602 <KIT>  
 A;Cross-references: UNIPROT:Q63921; EMBL:U18060; NID:G603051; PIDN:AAA85823.1; PID:G603051  
 R;Kitzler, J.; Hill, E.; Hardman, R.; Reddy, N.; Philpot, R.; Eling, T.E.  
 Arch. Biochem. Biophys. 316, 856-863, 1995  
 A;Title: Analysis and quantitation of splicing variants of the TPA-inducible PGHS-1 mRNA  
 A;Reference number: S69199; MUID:95168876; PMID:7864644  
 A;Accession: S69199

A;Molecule type: mRNA  
 A;Residues: 61-602 <KIT2>  
 A;Cross-references: EMBL:U18060  
 A;Note: only a part of the nucleic acid sequence is shown  
 C;Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 C;Keywords: alternative splicing  
 F;38-71/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLEFD 7  
 |||||  
 Db 358 FLQLKFD 364

Search completed: December 30, 2004, 20:52:20  
 Job time : 2.97229 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:29:59 ; Search time 10.5642 Seconds  
(without alignments)  
490.180 Million cell updates/sec

Title: US-10-017-327-5  
Perfect score: 44  
Sequence: 1 FLOLEFPAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02:  
1: uniprot\_sprot;  
2: uniprot\_trembl;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	84.1	124	Q94TV4	Q94tv4 hapalogaste
2	37	84.1	140	Q94TU9	Q94tu9 labidochiru
3	37	84.1	142	Q94TU8	Q94tu8 lepholithod
4	37	84.1	147	Q94TU6	Q94tu6 pagurus ber
5	37	84.1	155	Q94TV7	Q94tv7 cryptolitho
6	37	84.1	163	Q94TV6	Q94tv6 cryptolitho
7	37	84.1	166	Q94TV3	Q8wfy3 lepidopa ca
8	37	84.1	170	Q8WFX4	Q8wfy4 blepharipod
9	37	84.1	171	Q94TU3	Q94tu3 paralomis g
10	37	84.1	172	Q94TV7	Q94tv7 oedignathus
11	37	84.1	174	Q94TV2	Q94tv2 phyllolitho
12	37	84.1	178	Q94TV3	Q94tv3 hapalogaste
13	37	84.1	179	Q94TV1	Q94tv1 rhinolithod
14	37	84.1	180	Q94PLV2	Q6plv2 aegla scamo
15	37	84.1	180	Q94TV0	Q94tv0 lithodes sa
16	37	84.1	180	AAT08408	Aat08408 aegla sca
17	37	84.1	181	Q94TV2	Q94tv2 lithodes ae
18	37	84.1	183	Q94TV5	Q94tv5 glyptolitho
19	37	84.1	183	Q8WFZ3	Q8wzf3 raninolitho
20	37	84.1	184	Q8WFX8	Q8wfy8 munida quad
21	37	84.1	185	Q8WFX7	Q8wfy7 aegla urugu
22	37	84.1	188	Q94TV1	Q94tv1 lithodes ma
23	37	84.1	189	Q94PLS9	Q6pls9 aegla sp. k
24	37	84.1	189	Q94PLT0	Q6plt0 aegla sp. k
25	37	84.1	189	Q94PLT1	Q6plt1 aegla neuqu
26	37	84.1	189	Q94PLT3	Q6plt3 aegla inter
27	37	84.1	189	Q94PLT4	Q6plt4 aegla inter
28	37	84.1	189	Q94PLT5	Q6plt5 aegla inter
29	37	84.1	189	Q94PLT6	Q6plt6 aegla plate
30	37	84.1	189	Q94PLT7	Q6plt7 aegla plate
31	37	84.1	189	Q94PLT9	Q6plt9 aegla humah

32	37	84.1	189	2	Q6PLU0	Q6plu0 aegla sp. k
33	37	84.1	189	2	Q6PLU1	Q6plu1 aegla sp. k
34	37	84.1	189	2	Q6PLU2	Q6plu2 aegla ringu
35	37	84.1	189	2	Q6PLU4	Q6plu4 aegla sanlo
36	37	84.1	189	2	Q6PLU5	Q6plu5 aegla sanlo
37	37	84.1	189	2	Q6PLU6	Q6plu6 aegla jujuy
38	37	84.1	189	2	Q6PLU8	Q6plu8 aegla septe
39	37	84.1	189	2	Q6PLV0	Q6plv0 aegla scamo
40	37	84.1	189	2	Q6PLV3	Q6plv3 aegla obati
41	37	84.1	189	2	Q6PLV5	Q6plv5 aegla plate
42	37	84.1	189	2	Q6PLV6	Q6plv6 aegla plate
43	37	84.1	189	2	Q6PLV8	Q6plv8 aegla margi
44	37	84.1	189	2	Q6PLV9	Q6plv9 aegla strin
45	37	84.1	189	2	Q6PLW1	Q6plw1 aegla lepto

## ALIGNMENTS

RESULT 1  
Q94TV4  
ID Q94TV4 PRELIMINARY; PRT; 124 AA.  
AC Q94TV4;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Cytochrome oxidase subunit II (Fragment).  
OS Hapalogaster dentata.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
OC Lithodidae; Hapalogaster.  
OX NCBI\_TaxID=174385;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zaklan S.D., Cunningham C.W.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the binuclear center of the catalytic subunit 1 (By similarity).  
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome c + 2 H(2)O.  
CC -!- COPACTOR: Copper A (By similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).  
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family. EMBL; AF425366; AAL26524.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005739; C:mitochondrion; IEA.  
DR GO; GO:0005507; F:copper ion binding; IEA.  
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR008972; Cupredoxin.  
DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
DR Pfam; PF00116; COX2; 1.  
DR Pfam; PF02790; COX2\_TM; 1.  
DR PRINTS; PR01166; CYCOXIDASEII.  
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
FT NON\_TER 1  
FT NON\_TER 124 124  
SQ SEQUENCE 124 AA; 14660 MW; 57DC335EF2212FBI CRC64;

Query Match 84.1%; Score 37; DB 2; Length 124;  
Best Local Similarity 87.5%; Pred. No. 6.2;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLOLEFPA 8  
DB 93 FLOLEFDS 100



CC inner membrane (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425374; AAL26532.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CUA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.  
 KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 147 147  
 SQ SEQUENCE 147 AA; 17357 MW; 51BE16CED00DA004 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 147;  
 Best Local Similarity 87.5%; Pred. No. 7.4;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLEFDA 8  
 Db |||||:  
 97 FLQLEFDS 104  
 RESULT 5  
 Q94TV7  
 ID Q94TV7 PRELIMINARY; PRT; 155 AA.  
 AC Q94TV7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Cryptolithodes sitchensis.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lithodidae; Cryptolithodes.  
 OX NCBI\_TaxID=174327;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 CC 3 form the functional core of the enzyme complex. Subunit 2  
 CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome  
 CC c + 2 H(2)O.  
 CC -!- COFACTOR: Copper A (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425374; AAL26522.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CUA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.  
 KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 147 147  
 SQ SEQUENCE 147 AA; 17357 MW; 51BE16CED00DA004 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 147;  
 Best Local Similarity 87.5%; Pred. No. 7.4;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLEFDA 8  
 Db |||||:  
 97 FLQLEFDS 104  
 RESULT 5  
 Q94TV7  
 ID Q94TV7 PRELIMINARY; PRT; 155 AA.  
 AC Q94TV7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Cryptolithodes sitchensis.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lithodidae; Cryptolithodes.  
 OX NCBI\_TaxID=174327;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 CC 3 form the functional core of the enzyme complex. Subunit 2  
 CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome  
 CC c + 2 H(2)O.  
 CC -!- COFACTOR: Copper A (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425374; AAL26522.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CUA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.

KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 155 155  
 SQ SEQUENCE 155 AA; 17780 MW; 663801A37950E345 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 155;  
 Best Local Similarity 87.5%; Pred. No. 7.8;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLEFDA 8  
 Db |||||:  
 76 FLQLEFDS 83  
 RESULT 6  
 Q94TV6  
 ID Q94TV6 PRELIMINARY; PRT; 163 AA.  
 AC Q94TV6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Cryptolithodes typicus.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lithodidae; Cryptolithodes.  
 OX NCBI\_TaxID=174328;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 CC 3 form the functional core of the enzyme complex. Subunit 2  
 CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome  
 CC c + 2 H(2)O.  
 CC -!- COFACTOR: Copper A (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425364; AAL26522.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CUA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.  
 KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 163 163  
 SQ SEQUENCE 163 AA; 18747 MW; B569F4CF5A80FDA8 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 163;  
 Best Local Similarity 87.5%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLEFDA 8  
 Db |||||:  
 79 FLQLEFDS 86

```

RESULT 7
Q8WFY3 PRELIMINARY; PRT; 166 AA.
AC Q8WFY3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Lepidopa californica.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Hippoidea;
OC Albulidae; Lepidopa.
OX NCBI_TaxID=177228;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21884466; PubMed=11886621;
RA Morrison C.L., Harvey A.W., Lavery S., Tieu K., Huang Y.,
RA Cunningham C.W.;
RT "Mitochondrial gene rearrangements confirm the parallel evolution of
the crab-like form.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:345-350(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. Subunit 2
transfers the electrons from cytochrome c via its binuclear copper
A center to the bimetallic center of the catalytic subunit 1 (By
similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AF437626; AAL31589.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON TER 166
FT NON TER 166
SQ SEQUENCE 166 AA; 19030 MW; C207194A2159CDB2 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 8.4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFDA 8
Db 89 FQLEFDS 96

RESULT 8
Q8WFY4 PRELIMINARY; PRT; 170 AA.
AC Q8WFY4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Blepharipoda.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

```

```

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Hippoidea;
OC Albulidae; Blepharipoda.
OX NCBI_TaxID=177216;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21884466; PubMed=11886621;
RA Morrison C.L., Harvey A.W., Lavery S., Tieu K., Huang Y.,
RA Cunningham C.W.;
RT "Mitochondrial gene rearrangements confirm the parallel evolution of
the crab-like form.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:345-350(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. Subunit 2
transfers the electrons from cytochrome c via its binuclear copper
A center to the bimetallic center of the catalytic subunit 1 (By
similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AF437625; AAL31588.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON TER 170
FT NON TER 170
SQ SEQUENCE 170 AA; 19721 MW; D4F5895D1AE7CA83 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 170;
Best Local Similarity 87.5%; Pred. No. 8.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQLEFDA 8
Db 113 FQLEFDS 120

RESULT 9
Q94TU3 PRELIMINARY; PRT; 171 AA.
AC Q94TU3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Paralomis granulosa.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Paralomis.
OX NCBI_TaxID=174405;
RN [1]
RP SEQUENCE FROM N.A.
RX Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. Subunit 2
transfers the electrons from cytochrome c via its binuclear copper
A center to the bimetallic center of the catalytic subunit 1 (By
similarity).

```

```

CC similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- COPACTOR: Copper A (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL; AF425377; AAL26535.1; -.
CC DR GO:0016020; C:membrane; IEA.
CC DR GO:0005739; C:mitochondrion; IEA.
CC DR GO:0005507; F:copper ion binding; IEA.
CC DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC DR GO:0006118; P:electron transport; IEA.
CC DR InterPro; IPR001505; Copper CUA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR InterPro; IPR002429; Cyt_c_ox_2.
CC DR PFam; PF00116; COX2; 1.
CC DR PFam; PF02790; COX2_TM; 1.
CC DR PRINTS; PR01166; CYCOXIDASEII.
CC DR ProDom; PD000131; Copper CUA; 1.
CC Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 171
FT NON_TER 171 171
SQ SEQUENCE 171 AA; 19619 MW; D26EBDDDD906235 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 171;
Best Local Similarity 87.5%; Pred. No. 8.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQLEFDD 8
Db 94 FQLEFDS 101

RESULT 10
Q94TU7
ID Q94TU7 PRELIMINARY; PRT; 172 AA.
AC Q94TU7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Oedignathus inermis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Oedignathus.
OX NCBI_TaxID=6743;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- COPACTOR: Copper A (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL; AF425378; AAL26536.1; -.
CC DR GO:0016020; C:membrane; IEA.
CC DR GO:0005739; C:mitochondrion; IEA.
CC DR GO:0005507; F:copper ion binding; IEA.
CC DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC DR GO:0006118; P:electron transport; IEA.
CC DR InterPro; IPR001505; Copper CUA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR InterPro; IPR002429; Cyt_c_ox_2.
CC DR PFam; PF00116; COX2; 1.
CC DR PFam; PF02790; COX2_TM; 1.
CC DR PRINTS; PR01166; CYCOXIDASEII.
CC DR ProDom; PD000131; Copper CUA; 1.
CC Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 171
FT NON_TER 171 171
SQ SEQUENCE 171 AA; 19619 MW; D26EBDDDD906235 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 171;
Best Local Similarity 87.5%; Pred. No. 8.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQLEFDD 8
Db 94 FQLEFDS 101

RESULT 10
Q94TU7
ID Q94TU7 PRELIMINARY; PRT; 172 AA.
AC Q94TU7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Oedignathus inermis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Oedignathus.
OX NCBI_TaxID=6743;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- COPACTOR: Copper A (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL; AF425373; AAL26531.1; -.
CC DR GO:0016020; C:membrane; IEA.
CC DR GO:0005739; C:mitochondrion; IEA.
CC DR GO:0005507; F:copper ion binding; IEA.
CC DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC DR GO:0006118; P:electron transport; IEA.
CC DR InterPro; IPR001505; Copper CUA.
CC DR InterPro; IPR008972; Cupredoxin.

```

```

DR InterPro; IPR002429; Cyt_c_ox_2.
DR PFam; PF00116; COX2; 1.
DR PFam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper CUA; 1.
DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 172
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19918 MW; 8A397959C0867049 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 172;
Best Local Similarity 87.5%; Pred. No. 8.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQLEFDD 8
Db 97 FQLEFDS 104

RESULT 11
Q94TU2
ID Q94TU2 PRELIMINARY; PRT; 174 AA.
AC Q94TU2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Phylloolithodes papillosus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Phylloolithodes.
OX NCBI_TaxID=174407;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- COPACTOR: Copper A (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL; AF425378; AAL26536.1; -.
CC DR GO:0016020; C:membrane; IEA.
CC DR GO:0005739; C:mitochondrion; IEA.
CC DR GO:0005507; F:copper ion binding; IEA.
CC DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC DR GO:0006118; P:electron transport; IEA.
CC DR InterPro; IPR001505; Copper CUA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR InterPro; IPR002429; Cyt_c_ox_2.
CC DR PFam; PF00116; COX2; 1.
CC DR PFam; PF02790; COX2_TM; 1.
CC DR PRINTS; PR01166; CYCOXIDASEII.
CC DR ProDom; PD000131; Copper CUA; 1.
CC Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 174
FT NON_TER 174 174
SQ SEQUENCE 174 AA; 19948 MW; EB262FBEF0CED389 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 174;
Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FLQLEFDA 8
Db 90 FLQLEFDS 97

RESULT 12
Q94TV3 PRELIMINARY; PRT; 178 AA.
AC Q94TV3,
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Cytochrome oxidase subunit II (Fragment).
OS Haplogaster mertensii.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Haplogaster.
OX NCBI_TaxID=174386;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AF425379; AAL26525.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 178
SQ SEQUENCE 178 AA; 20305 MW; D21D40F3CE64D94F CRC64;

Query Match 84.1%; Score 37; DB 2; Length 178;
Best Local Similarity 87.5%; Pred. No. 9.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLEFDA 8
Db 90 FLQLEFDS 97

RESULT 13
Q94TV1 PRELIMINARY; PRT; 179 AA.
AC Q94TV1,
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Cytochrome oxidase subunit II (Fragment).
OS Rhinolithodes wosnesenskii.
OG Mitochondrion.

```

```

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Rhinolithodes.
OX NCBI_TaxID=174409;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AF425379; AAL26537.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 179
SQ SEQUENCE 179 AA; 20598 MW; 18058F40B6B1B2A8 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 179;
Best Local Similarity 87.5%; Pred. No. 9.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLQLEFDA 8
Db 93 FLQLEFDS 100

RESULT 14
Q6PLV2 PRELIMINARY; PRT; 180 AA.
AC Q6PLV2,
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE Cytochrome oxidase subunit II (Fragment).
OS Aegla scamosa.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegidae; Aegla.
OX NCBI_TaxID=273843;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RL "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2

```

transfers the electrons from cytochrome c via its binuclear copper  
A center to the bimetallic center of the catalytic subunit 1 (By  
similarity).  
-|- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
c + 2 H(2)O.  
-|- COPACTOR: Copper A (By similarity).  
-|- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
inner membrane (By similarity).  
-|- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
DR EMBL: AY595772; AAT08408.1; -.  
DR GO: GO:0005739; C:mitochondrion; IEA.  
DR InterPro: IPR001505; Copper\_CuA.  
DR InterPro: IPR008972; Cupredoxin.  
DR Pfam: PF00116; COX2; 1.  
DR Pfam: PF02790; COX2\_TM; 1.  
DR PRINTS: PR01166; CYCOXIDASEII.  
DR ProDom: PD000131; Copper\_CuA; 1.  
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
FT NON\_TER 1 180  
SQ SEQUENCE 180 AA; 20778 MW; C827D7E7F4CA2C0B CRC64;  
  
Query Match 84.1%; Score 37; DB 2; Length 180;  
Best Local Similarity 87.5%; Pred. No. 9.2;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FLOLEFDA 8  
| | | | | | | |  
Db 92 FLOLEFDS 99  
  
RESULT 15  
Q94TV0  
ID Q94TV0 PRELIMINARY; PRT; 180 AA.  
AC Q94TV0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cytochrome oxidase subunit II (Fragment).  
OS Lithodes santolla.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
OC Lithodidae; Lithodes.  
OX NCBI\_TaxID=174394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zaklan S.D., Cunningham C.W.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -|- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
chain that catalyzes the reduction of oxygen to water. Subunits 1-  
3 form the functional core of the enzyme complex. Subunit 2  
transfers the electrons from cytochrome c via its binuclear copper  
A center to the bimetallic center of the catalytic subunit 1 (By  
similarity).  
CC -|- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
c + 2 H(2)O.  
CC -|- COPACTOR: Copper A (By similarity).  
CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
inner membrane (By similarity).  
CC -|- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
DR EMBL: AF425370; AAL26528.1; -.  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005739; C:mitochondrion; IEA.  
DR GO: GO:0005507; F:copper ion binding; IEA.  
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.  
DR GO: GO:0006118; P:electron transport; IEA.  
DR InterPro: IPR001505; Copper\_CuA.  
DR InterPro: IPR008972; Cupredoxin.  
DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
DR Pfam: PF00116; COX2; 1.

DR Pfam: PF02790; COX2\_TM; 1.  
DR PRINTS: PR01166; CYCOXIDASEII.  
DR ProDom: PD000131; Copper\_CuA; 1.  
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
FT NON\_TER 1 180  
SQ SEQUENCE 180 AA; 20816 MW; F4ED7795D4F090F1 CRC64;  
  
Query Match 84.1%; Score 37; DB 2; Length 180;  
Best Local Similarity 87.5%; Pred. No. 9.2;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FLOLEFDA 8  
| | | | | | | |  
Db 97 FLOLEFDS 104  
  
Search completed: December 30, 2004, 20:50:50  
Job time : 11.5642 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:34:10 ; Search time 2.06297 Seconds  
(without alignments)  
289.321 Million cell updates/sec

Title: US-10-017-327-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	78.0	585	US-09-489-039A-13868	Sequence 13868, A
2	36	72.0	373	US-09-252-991A-24034	Sequence 24034, A
3	36	72.0	700	US-09-266-225D-6	Sequence 6, Appli
4	35	70.0	543	US-09-540-236-2405	Sequence 2405, Ap
5	35	70.0	599	US-09-228-986-117	Sequence 117, App
6	35	70.0	599	US-10-101-464A-117	Sequence 117, App
7	35	70.0	816	US-10-101-464A-827	Sequence 827, App
8	35	70.0	951	US-09-328-352-4456	Sequence 4456, Ap
9	35	70.0	1240	US-10-101-464A-976	Sequence 976, App
10	34	68.0	262	US-09-252-991A-29018	Sequence 29018, A
11	33	66.0	86	US-09-270-767-60013	Sequence 60013, A
12	33	66.0	309	US-08-952-736A-10	Sequence 10, Appl
13	33	66.0	338	US-09-198-452A-1050	Sequence 1050, Ap
14	33	66.0	363	US-09-328-352-5961	Sequence 5961, Ap
15	33	66.0	516	US-08-676-166A-2	Sequence 2, Appli
16	32	64.0	68	US-09-513-999C-4712	Sequence 4712, Ap
17	32	64.0	120	US-09-489-039A-8587	Sequence 8587, Ap
18	32	64.0	231	US-09-134-001C-4094	Sequence 4094, Ap
19	32	64.0	279	US-09-328-352-5581	Sequence 5581, Ap
20	32	64.0	308	US-09-328-352-7093	Sequence 7093, Ap
21	32	64.0	323	US-09-270-767-31938	Sequence 31938, A
22	32	64.0	328	US-09-583-110-4999	Sequence 4999, Ap
23	32	64.0	350	US-08-415-751-43	Sequence 43, Appl
24	32	64.0	419	US-09-489-039A-12789	Sequence 12789, A
25	32	64.0	480	US-09-107-532A-6160	Sequence 6160, Ap
26	32	64.0	481	US-08-617-785-8	Sequence 8, Appli
27	32	64.0	481	US-09-817-464-8	Sequence 8, Appli

28	32	64.0	511	1	US-08-220-151-17	Sequence 17, Appl
29	32	64.0	511	1	US-08-413-118-17	Sequence 17, Appl
30	32	64.0	511	3	US-08-473-446-17	Sequence 17, Appl
31	32	64.0	575	4	US-09-248-796A-17643	Sequence 17643, A
32	32	64.0	754	4	US-09-252-991A-33133	Sequence 33133, A
33	32	64.0	828	4	US-10-101-464A-934	Sequence 934, App
34	32	64.0	867	3	US-08-617-785-4	Sequence 4, Appli
35	32	64.0	867	4	US-09-817-464-4	Sequence 4, Appli
36	32	64.0	915	1	US-08-453-862-2	Sequence 2, Appli
37	32	64.0	915	2	US-08-452-734A-2	Sequence 2, Appli
38	32	64.0	915	3	US-08-617-785-12	Sequence 12, Appl
39	32	64.0	915	3	US-08-176-401B-2	Sequence 2, Appli
40	32	64.0	915	4	US-09-817-464-12	Sequence 12, Appl
41	32	64.0	915	5	PCT-US94-14989-2	Sequence 2, Appli
42	32	64.0	922	3	US-08-617-785-14	Sequence 14, Appl
43	32	64.0	922	4	US-09-817-464-14	Sequence 14, Appl
44	31	62.0	63	4	US-09-248-796A-21418	Sequence 21418, A
45	31	62.0	65	4	US-09-513-999C-5288	Sequence 5288, Ap

ALIGNMENTS

RESULT 1  
US-09-489-039A-13868  
; Sequence 13868, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13868  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13868

Query Match 78.0%; Score 39; DB 4; Length 585;  
Best Local Similarity 85.7%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
Db 273 WFEIDVV 279  
|||||

RESULT 2  
US-09-252-991A-24034  
; Sequence 24034, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24034  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24034

```
Query Match      72.0%; Score 36; DB 4; Length 373;
Best Local Similarity 55.6%; Pred. No. 85;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
Db 307 YLWVELDIV 315

RESULT 3
US-09-266-225D-6
; Sequence 6, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-6

Query Match      72.0%; Score 36; DB 4; Length 700;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDI 8
Db 612 FLWFEIDM 619

RESULT 4
US-09-540-236-2405
; Sequence 2405, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2405
; LENGTH: 543
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2405

Query Match      70.0%; Score 35; DB 4; Length 543;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
Db 308 FFWTEIDLI 316

RESULT 5
US-09-228-986-117
; Sequence 117, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2

Query Match      70.0%; Score 35; DB 3; Length 599;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFEID 7
Db 33 LWFEVD 38

RESULT 6
US-10-101-464A-117
; Sequence 117, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-117

Query Match      70.0%; Score 35; DB 4; Length 599;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFEID 7
Db 33 LWFEVD 38

RESULT 7
US-10-101-464A-827
; Sequence 827, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
```

; CURRENT APPLICATION NUMBER: US/10/101,464A  
 ; CURRENT FILING DATE: 2002-03-18  
 ; PRIOR APPLICATION NUMBER: 09/704,302  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: 09/228,986  
 ; PRIOR FILING DATE: 1999-01-12  
 ; PRIOR APPLICATION NUMBER: 60/162,866  
 ; PRIOR FILING DATE: 1999-11-01  
 ; PRIOR APPLICATION NUMBER: PCT/US00/00724  
 ; PRIOR FILING DATE: 2000-01-11  
 ; NUMBER OF SEQ ID NOS: 989  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 827  
 ; LENGTH: 816  
 ; TYPE: PRT  
 ; ORGANISM: Eucalyptus grandis  
 US-10-101-464A-827

Query Match 70.0%; Score 35; DB 4; Length 816;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEID 7  
 ||||:|  
 Db 250 LWFEVD 255

## RESULT 8

US-09-328-352-4456  
 ; Sequence 4456, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 4456  
 ; LENGTH: 951  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-4456

Query Match 70.0%; Score 35; DB 4; Length 951;  
 Best Local Similarity 50.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEIDIV 9  
 ||||:|  
 Db 801 WFKVDVV 808

## RESULT 9

US-10-101-464A-976  
 ; Sequence 976, Application US/10101464A  
 ; Patent No. 6768041  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Strabala, Timothy  
 ; APPLICANT: Nieuwenhuizen, Nicolaas  
 ; APPLICANT: Higgins, Colleen M.  
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
 ; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
 ; FILE REFERENCE: 11000.1020c2  
 ; CURRENT APPLICATION NUMBER: US/10/101,464A  
 ; CURRENT FILING DATE: 2002-03-18  
 ; PRIOR APPLICATION NUMBER: 09/704,302  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: 09/228,986  
 ; PRIOR FILING DATE: 1999-01-12  
 ; PRIOR APPLICATION NUMBER: 60/162,866  
 ; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724  
 ; PRIOR FILING DATE: 2000-01-11  
 ; NUMBER OF SEQ ID NOS: 989  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 976  
 ; LENGTH: 1240  
 ; TYPE: PRT  
 ; ORGANISM: Eucalyptus grandis  
 US-10-101-464A-976

Query Match 70.0%; Score 35; DB 4; Length 1240;  
 Best Local Similarity 83.3%; Pred. No. 4.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEID 7  
 ||||:|  
 Db 674 LWFEVD 679

## RESULT 10

US-09-252-991A-29018  
 ; Sequence 29018, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 29018  
 ; LENGTH: 262  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-29018

Query Match 68.0%; Score 34; DB 4; Length 262;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEID 7  
 ||||:|  
 Db 110 LWFEVD 115

## RESULT 11

US-09-270-767-60013  
 ; Sequence 60013, Application US/09270767  
 ; Patent No. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Homburger et al.  
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 ; FILE REFERENCE: File Reference: 7326-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; CURRENT FILING DATE: 1999-03-17  
 ; NUMBER OF SEQ ID NOS: 62517  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 60013  
 ; LENGTH: 86  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 ; FEATURE:  
 ; OTHER INFORMATION: Xaa means any amino acid  
 US-09-270-767-60013

Query Match 66.0%; Score 33; DB 4; Length 86;  
 Best Local Similarity 57.1%; Pred. No. 62;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
QY      1 FLWFEID 7
      78 FIWFDFD 84

US-08-952-736A-10
; Sequence 10, Application US/08952736A
; Patent No. 6320026
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cell Growth Inhibitory Factor
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,736A
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-952-736A-10

Query Match      66.0%; Score 33; DB 3; Length 309;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLWFEID 7
      29 FKWFQID 35

US-09-198-452A-1050
; Sequence 1050, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1050
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1050

Query Match      66.0%; Score 33; DB 4; Length 338;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLWFEID 7
      115 FIWFEXD 121

US-09-328-352-5961
; Sequence 5961, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5961
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5961

Query Match      66.0%; Score 33; DB 4; Length 363;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLWFEID 7
      133 FIWFQID 139

US-08-676-166A-2
; Sequence 2, Application US/08676166A
; Patent No. 5955270
; GENERAL INFORMATION:
; APPLICANT: Radford, Alan
; APPLICANT: Parish, John H.
; TITLE OF INVENTION: EXPLOITATION OF THE CELLULASE COMPLEX OF
; TITLE OF INVENTION: NEUROSPORA
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,166A
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-676-166A-2

Query Match      66.0%; Score 33; DB 2; Length 516;
Best Local Similarity 57.1%; Pred. No. 4.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LWFEIDI 8
      139 LWFDVDM 145

Search completed: December 30, 2004, 20:53:56
```

Job time : 3.06297 secs

---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:35:11 ; Search time 8.81864 Seconds  
(without alignments)  
367.126 Million cell updates/sec

Title: US-10-017-327-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues  
Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	11 US-09-870-216C-7	Sequence 7, Appli
2	50	100.0	9	13 US-10-017-327-7	Sequence 7, Appli
3	39	78.0	94	17 US-10-425-115-206553	Sequence 206553,
4	39	78.0	1717	14 US-10-126-339-2	Sequence 2, Appli
5	39	78.0	1717	14 US-10-126-148-2	Sequence 2, Appli
6	39	78.0	1717	17 US-10-798-058-8	Sequence 8, Appli
7	38	76.0	112	15 US-10-424-599-183462	Sequence 183462,
8	37	74.0	529	14 US-10-369-493-18446	Sequence 18446, A
9	36	72.0	58	15 US-10-424-599-144215	Sequence 144215,
10	36	72.0	402	14 US-10-369-493-20052	Sequence 20052, A
11	36	72.0	700	16 US-10-408-765A-2042	Sequence 2042, Ap
12	36	72.0	1759	17 US-10-798-058-2	Sequence 2, Appli
13	35	70.0	37	15 US-10-424-599-273756	Sequence 273756,

14	35	70.0	210	15	US-10-424-599-236699	Sequence 236699,
15	35	70.0	217	15	US-10-041-615-63	Sequence 63, Appli
16	35	70.0	217	15	US-10-072-012-557	Sequence 557, App
17	35	70.0	232	14	US-10-017-161-1002	Sequence 1002, Ap
18	35	70.0	232	14	US-10-292-798-852	Sequence 852, App
19	35	70.0	315	15	US-10-041-615-16	Sequence 16, Appli
20	35	70.0	315	15	US-10-072-012-192	Sequence 192, App
21	35	70.0	326	14	US-10-017-161-460	Sequence 460, App
22	35	70.0	326	14	US-10-017-161-506	Sequence 506, App
23	35	70.0	326	14	US-10-387-629-6	Sequence 6, Appli
24	35	70.0	326	14	US-10-292-798-412	Sequence 412, App
25	35	70.0	326	14	US-10-292-798-446	Sequence 446, App
26	35	70.0	326	15	US-10-343-650A-640	Sequence 640, App
27	35	70.0	599	14	US-10-101-464A-117	Sequence 117, App
28	35	70.0	720	15	US-10-425-114-72890	Sequence 72890, A
29	35	70.0	816	14	US-10-101-464A-827	Sequence 827, App
30	35	70.0	1240	14	US-10-101-464A-976	Sequence 976, App
31	34	68.0	44	17	US-10-425-115-227895	Sequence 227895,
32	34	68.0	51	9	US-09-798-029-21	Sequence 21, Appli
33	34	68.0	61	16	US-10-437-963-147599	Sequence 147599,
34	34	68.0	62	17	US-10-425-115-363307	Sequence 363307,
35	34	68.0	75	15	US-10-424-599-218578	Sequence 218578,
36	34	68.0	77	15	US-10-424-599-152779	Sequence 152779,
37	34	68.0	175	15	US-10-424-599-264160	Sequence 264160,
38	34	68.0	175	15	US-10-424-599-264161	Sequence 264161,
39	34	68.0	178	15	US-10-424-599-261009	Sequence 261009,
40	34	68.0	215	15	US-10-424-599-264158	Sequence 264158,
41	34	68.0	296	15	US-10-282-122A-53995	Sequence 53995, A
42	34	68.0	378	16	US-10-159-257A-176	Sequence 176, App
43	34	68.0	393	15	US-10-424-599-177498	Sequence 177498,
44	34	68.0	416	15	US-10-282-122A-48716	Sequence 48716, A
45	34	68.0	422	15	US-10-471-243-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-09-870-216C-7  
; Sequence 7, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870.216C  
; PRIORITY FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-216C-7

Query Match 100.0%; Score 50; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
| | | | |  
Db 1 FLWFEIDIV 9

RESULT 2  
US-10-017-327-7  
; Sequence 7, Application US/10017327

```
; Publication No. US20020155471A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
; FILE REFERENCE: G2 2101.20
; CURRENT APPLICATION NUMBER: US/10/017,327
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-327-7

Query Match      100.0%; Score 50; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
Db 1 FLWFEIDIV 9

RESULT 3
US-10-425-115-206553
; Sequence 206553, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 363326
; SEQ ID NO 206553
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(94)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_119961C.1.pep
US-10-425-115-206553

Query Match      78.0%; Score 39; DB 17; Length 94;
Best Local Similarity 55.6%; Pred. No. 27;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
Db 26 FLWVELDYI 34

RESULT 4
US-10-126-339-2
; Sequence 2, Application US/10126339
; Publication No. US20030166891A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/10/126,339
; CURRENT FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
```

```
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-10-126-339-2

Query Match      78.0%; Score 39; DB 14; Length 1717;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7
Db 650 YLMFEID 656

RESULT 5
US-10-126-148-2
; Sequence 2, Application US/10126148
; Publication No. US20030170787A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: No. US20030170787A1 Bt Toxin Receptors From
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/10/126,148
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-10-126-148-2

Query Match      78.0%; Score 39; DB 14; Length 1717;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7
Db 650 YLMFEID 656

RESULT 6
US-10-798-058-8
; Sequence 8, Application US/10798058
; Publication No. US20040209329A1
; GENERAL INFORMATION:
; APPLICANT: Mathis, John P.
; TITLE OF INVENTION: Novel Bt Toxin Receptors and Methods of
; FILE REFERENCE: 35718/274644
; CURRENT APPLICATION NUMBER: US/10/798,058
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: US 60/455,085
; PRIOR FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
```



US-10-798-058-8

Query Match 78.0%; Score 39; DB 17; Length 1717;  
Best Local Similarity 85.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7  
:|||||  
Db 650 YLWFEID 656

## RESULT 7

US-10-424-599-183462  
; Sequence 183462, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 183462  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(112)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13667C.1.pap  
US-10-424-599-183462

Query Match 76.0%; Score 38; DB 15; Length 112;  
Best Local Similarity 77.8%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
|||||  
Db 39 FLWFNDKV 47

## RESULT 8

US-10-369-493-18446  
; Sequence 18446, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18446  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Lactococcus lactis  
US-10-369-493-18446

Query Match 74.0%; Score 37; DB 14; Length 529;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
|||||  
Db 49 FLWLAIDII 57

## RESULT 9

US-10-424-599-144215  
; Sequence 144215, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 144215  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101239C.1.pap  
US-10-424-599-144215

Query Match 72.0%; Score 36; DB 15; Length 58;  
Best Local Similarity 75.0%; Pred. No. 58;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LWFEIDIV 9  
|||||  
Db 45 LWFEVIDII 52

## RESULT 10

US-10-369-493-20052  
; Sequence 20052, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 20052  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: No. US20030233675Altoc punctiforme  
US-10-369-493-20052

Query Match 72.0%; Score 36; DB 14; Length 402;  
Best Local Similarity 71.4%; Pred. No. 3.6e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEIDIV 9  
|||||  
Db 245 WFEVDIL 251

## RESULT 11

US-10-408-765A-2042  
; Sequence 2042, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warrock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2042  
; LENGTH: 700  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-2042

Query Match 72.0%; Score 36; DB 16; Length 700;  
Best Local Similarity 75.0%; Pred. No. 6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDI 8  
| | | | | :  
DB 612 FLWFENDM 619

## RESULT 12

US-10-798-058-2  
; Sequence 2, Application US/10798058  
; Publication No. US20040209329A1  
; GENERAL INFORMATION:  
; APPLICANT: Mathis, John P.  
; TITLE OF INVENTION: Novel Bt Toxin Receptors and Methods of  
; FILE REFERENCE: 35718/274644  
; CURRENT APPLICATION NUMBER: US/10/798,058  
; CURRENT FILING DATE: 2004-03-11  
; PRIOR APPLICATION NUMBER: US 60/455,085  
; PRIOR FILING DATE: 2003-03-14  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1759  
; TYPE: PRT  
; ORGANISM: Agrotis ipsilon  
US-10-798-058-2

Query Match 72.0%; Score 36; DB 17; Length 1759;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFEID 7  
| | | | | :  
DB 668 LWFEID 673

## RESULT 13

US-10-424-599-273756  
; Sequence 273756, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 273756  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_892222C.1.pap  
US-10-424-599-273756

Query Match 70.0%; Score 35; DB 15; Length 37;  
Best Local Similarity 71.4%; Pred. No. 57;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7  
| | | | | :  
DB 13 FLWFKMD 19

## RESULT 14

US-10-424-599-236699  
; Sequence 236699, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 236699  
; LENGTH: 210  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(210)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55766C.1.pap  
US-10-424-599-236699

Query Match 70.0%; Score 35; DB 15; Length 210;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEID 7  
| | | | | :  
DB 55 FWFWDID 61

## RESULT 15

US-10-041-615-63  
; Sequence 63, Application US/10041615  
; Publication No. US20040014038A1  
; GENERAL INFORMATION:  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Padigar, Muralidhara  
; TITLE OF INVENTION: No. US20040014038A1el GPCR-Like Proteins and Nucleic Acids Encodir  
; FILE REFERENCE: 21402-233-061  
; CURRENT APPLICATION NUMBER: US/10/041,615  
; CURRENT FILING DATE: 2003-01-29

```
; PRIOR APPLICATION NUMBER: 60/259,552
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/260,544
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/277,405
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 63
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-615-63

Query Match      70.0%; Score 35; DB 15; Length 217;
Best Local Similarity 77.8%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLWFEIDIV 9
      |||||
Db      86 FLWFLIPIV 94

Search completed: December 30, 2004, 21:00:33
Job time : 9.81864 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:25:59 : Search time 10.4962 Seconds  
(without alignments)  
307.593 Million cell updates/sec

**Title:** US-10-017-327-7  
**Perfect score:** 50  
**Sequence:** 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :
1: Genesexp13Sep04:*
2: Genesexp13Sep04:*
3: Genesexp19Sep04:*
4: Genesexp2000s:*
5: Genesexp2001s:*
6: Genesexp2002s:*
7: Genesexp2003a:*
8: Genesexp2003b:*
9: Genesexp2004s:*
```

## SUMMARIES

Result No.	Query No.	Score	Query			DB	ID	Description
			Match	Length	%			
1	50	100.0	9	5	AB0808363	Ab0808363	Synthetic	
2	50	100.0	9	7	ABR82215	AbR82215	Human ant	
3	50	100.0	352	5	AB0808368	Ab0808368	Human can	
4	39	78.0	585	7	AB067351	Ab067351	Klebsiell	
5	39	78.0	587	6	AAC030104	Aac030104	Yersinia	
6	39	78.0	1717	4	AB020498	Ab020498	Oestrinia	
7	38	76.0	589	6	ABM69298	Abm69298	Photorhab	
8	37	74.0	150	5	ABM49338	Abm49338	Listeria	
9	37	74.0	273	5	ABP64842	Abp64842	Human pro	
10	37	74.0	528	5	AB554865	Ab554865	Lactococc	
11	37	74.0	587	6	ABM70700	Abm70700	Photorhab	
12	36	72.0	269	5	AB853431	Ab853431	Lactococc	
13	36	72.0	280	4	ABG01152	Abg01152	Novel hum	
14	36	72.0	373	7	AB075288	Ab075288	Pseudomon	
15	36	72.0	410	4	ABG01159	Abg01159	Novel hum	
16	36	72.0	591	6	ABM68497	Abm68497	Photorhab	
17	36	72.0	652	6	AB014689	Ab014689	Novel hum	
18	36	72.0	700	6	AB014690	Ab014690	Novel hum	
19	36	72.0	700	7	ADJ14125	Adj14125	Human src	
20	36	72.0	700	7	ADP70236	Adp70236	Human hea	
21	36	72.0	1097	4	ABG01151	Abg01151	Novel hum	
22	35	70.0	217	5	AD117021	Ad117021	Human NOV	
23	35	70.0	232	7	ADC86399	Adc86399	Human GPC	
24	35	70.0	307	4	AAG71459	Aag71459	Human olf	
25	35	70.0	315	4	AAG72215	Aag72215	Human olf	

## ALIGNMENTS

RESULT 1	
ABB08363	
ID	ABB08363 standard; protein; 9 AA.
XX	
XX	ABB08363;
XX	
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Synthetic epitope 3 of human cancer antigen eIF3.
XX	
KW	Human; melanoma antigen eukaryotic initiation factor 3; eIF3;
KW	ovarian cancer; MHC; cytostatic; immunomodulator; immune effe
KW	anti-cancer; vaccine.
XX	
XX	
OS	Homo sapiens.
XX	
XX	
PH	Key
FT	Domain
FT	/note= "HLA-2 binding residue"
FT	1
FT	Domain
FT	/note= "HLA-2 binding residue"
FT	2
FT	Domain
FT	/note= "HLA-2 binding residue"
FT	3..8
FT	Domain
FT	/note= "T-cell receptor (TCR) binding domain"
FT	9
FT	Domain
FT	/note= "HLA-2 binding residue"
FT	9

XX The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory  
CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.  
CC Compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterised by expression of antigen eIF3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detecting genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents synthetic epitope 3 of human cancer antigen eIF3  
XX  
XX Sequence 9 AA;

```

Query Watch      100.0%; Score 50; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLWFEIDIV 9
        |||||
Db       1 FLWFEIDIV 9

RESULT 2
ABR82215
ID  ABR82215 standard; peptide; 9 AA.
XX
AC  ABR82215;
XX
DT  13-OCT-2003 (first entry)
XX
DE  Human antigen eIF3 derived compound 3.
XX
KW  Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;
KW  cytostatic; gene therapy; human; antigen.

```

CC	condition or susceptibility to a neoplastic condition based on the amount
CC	of expression of the eIF3 protein. The methods, compounds and kits are
CC	useful in therapeutics, diagnostic and screening methods for human cancer
CC	and related malignancies, e.g. ovarian, breast, lung, colon, prostate,
CC	pancreatic or gastrointestinal cancer, or melanoma. Sequences ABR82213-16
CC	represent compounds derived from the human antigen eIF3
XX	
SQ	Sequence 9 AA;
<p>Query Match 100.0%; Score 50; DB 7; Length 9;            Best Local Similarity 100.0%; Pred. No. 1.7e+06;            Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0</p>	
QY	1 FLWFEIDIV 9 
Db	1 FLWFEIDIV 9
<p>RESULT 3            ABB08368            ID ABB08368 standard; protein; 352 AA.            AC ABB08368;            XX            DT 07-MAY-2002 (first entry)            XX            DE Human cancer antigen eIF3 variant 3 amino acid sequence.            XX            KW Human; melanoma antigen eukaryotic initiation factor 3; eIF3;            KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;            KW anti-cancer; vaccine.            XX            OS Homo sapiens.            XX</p>	
FH	Key Location/Qualifiers
FT	Misc-difference 242 /note= "wild-type Asn is replaced by Phe"
FT	Domain 242 /note= "HLA-2 binding residue"
FT	Domain 243 /note= "HLA-2 binding residue"
FT	Domain 244.249 /note= "T-cell receptor (TCR) binding domain"
FT	Misc-difference 244 /note= "wild-type Gln is replaced by Trp"
FT	Misc-difference 245 /note= "wild-type Leu is replaced by Phe"
FT	Misc-difference 246 /note= "wild-type Leu is replaced by Glu"
FT	Misc-difference 247 /note= "wild-type Met is replaced by Ile"
FT	Misc-difference 249 /note= "wild-type Arg is replaced by Ile"
FT	Domain 250 /note= "HLA-2 binding residue"
XX	WO200192307-A2.
XX	06-DEC-2001.
XX	30-MAY-2001; 2001WO-US017456.
XX	31-MAY-2000; 2000US-0209391P.
XX	17-AUG-2000; 2000US-0226258P.
XX	20-DEC-2000; 2000US-0257008P.
XX	(GENZ ) GENZYME CORP.
XX	Nicolette CA;
XX	WPI; 2002-139606/18.
XX	New therapeutic compounds useful against human ovarian cancer. for

RESULT 3	
ABB08368	
ID	ABB08368 standard; protein; 352 AA.
XX	
AC	ABB08368;
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Human cancer antigen eIF3 variant 3 amino acid sequence.
XX	
KW	Human; melanoma antigen eukaryotic initiation factor 3; eIF3;
KW	ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;
KW	anti-cancer; vaccine.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 242
FT	/note= "wild-type Asn is replaced by Phe"
FT	242
FT	/note= "HLA-2 binding residue"
FT	243
FT	/note= "HLA-2 binding residue"
FT	244..249
FT	/note= "T-cell receptor (TCR) binding domain"
FT	244
FT	/note= "wild-type Gln is replaced by Trp"
FT	245
FT	/note= "wild-type Leu is replaced by Phe"
FT	246
FT	/note= "wild-type Leu is replaced by Glu"
FT	247
FT	/note= "wild-type Met is replaced by Ile"
FT	249
FT	/note= "wild-type Arg is replaced by Ile"
FT	250
FT	/note= "HLA-2 binding residue"
XX	
PN	WO200192307-A2.
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001WO-US017456.
XX	
PR	31-MAY-2000; 2000US-0209391P.
PR	17-AUG-2000; 2000US-0226258P.
PR	20-DEC-2000; 2000US-0257008P.
XX	
PA	(GENZ ) GENZYME CORP.
XX	
PI	Nicolette CA;
XX	
DR	WPI; 2002-139606/18.
XX	
PT	New therapeutic compounds useful against human ovarian cancer. for

PT modulating immune response in a subject, and for generating antibodies  
 XX that specifically recognize and bind to these molecules.  
 PS Claim 8; Page; 68pp; English.  
 XX  
 CC The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen E1F3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human cancer antigen e1f3 variant 3 amino acid sequence.  
 CC Note: This sequence is not present in the specification, but may be  
 CC created from the sequence of the wild-type human cancer antigen e1f3  
 CC sequence given in ABB08360  
 XX  
 SQ Sequence 352 AA;  
 Query Match 100.0%; Score 50; DB 5; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 0.94; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;  
 QY 1 FLWFEIDIV 9  
 DB 242 FLWFEIDIV 250  
 RESULT 4  
 ABO67351  
 ID ABO67351 standard; protein; 585 AA.  
 XX  
 AC ABO67351;  
 XX  
 XX 29-JUL-2004 (first entry)  
 XX  
 DE Klebsiella pneumoniae polypeptide seqid 13868.  
 XX  
 KW Recombinant expression vector; transcription regulatory element;  
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.  
 XX  
 OS Klebsiella pneumoniae.  
 XX  
 PN US6610836-B1.  
 XX  
 XX 26-AUG-2003.  
 XX  
 XX 27-JAN-2000; 2000US-00489039.  
 XX  
 XX 29-JAN-1999; 99US-0117747P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton GL, Osborne M;  
 XX  
 XX WPI; 2003-895346/82.  
 DR N-PSDB; ABD00922.  
 XX  
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
 PT preparing a vaccine composition against Klebsiella pneumoniae.  
 XX  
 PS Disclosure; SEQ ID NO 13868; 932pp; English.  
 XX  
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella

CC pneumoniae polypeptide. Also described are: a recombinant expression  
 CC vector comprising the nucleic acid, operably linked to a transcription  
 CC regulatory element; and a cell comprising the recombinant expression  
 CC vector. The nucleic acid is useful for preparing a vaccine composition  
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a  
 CC Klebsiella pneumoniae polypeptide of the invention  
 XX  
 SQ Sequence 585 AA;  
 Query Match 78.0%; Score 39; DB 7; Length 585;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 1;  
 QY 3 WFEIDIV 9  
 DB 273 WFEIDIV 279  
 RESULT 5  
 AAO30104  
 ID AAO30104 standard; protein; 587 AA.  
 XX  
 AC AAO30104;  
 XX  
 DT 03-SEP-2003 (first entry)  
 XX  
 DE Yersinia pestis 2G5 mutant protein.  
 XX  
 KW Infection; tuberculosis; veterinary treatment; antimicrobial drug;  
 KW vaccine; mutant; mutein.  
 OS Yersinia pestis.  
 OS Synthetic.  
 XX  
 PN WO200304047-A2.  
 XX  
 PD 30-MAY-2003.  
 XX  
 PF 18-NOV-2002; 2002WO-GB005212.  
 XX  
 PR 19-NOV-2001; 2001GB-00027657.  
 XX  
 PA (UNLO ) LONDON SCHOOL HYGIENE & TROPICAL MEDICIN.  
 XX  
 PI Karlyshev AV, Wren BW;  
 XX  
 DR WPI; 2003-457588/43.  
 DR N-PSDB; AAL60720.  
 XX  
 PT New virulence peptides encoded by a gene of Yersinia pseudotuberculosis,  
 PT useful for treating or preventing a condition associated with infection  
 PT by Yersinia, e.g. tuberculosis, particularly in veterinary treatment.  
 XX  
 PS Claim 4; Page 96-99; 130pp; English.  
 XX  
 XX The present invention relates to virulence genes of Yersinia species and  
 CC proteins encoded by such genes. Sequences of the invention are useful in  
 CC the manufacture of a medicament for treating or preventing a condition  
 CC associated with infection by Yersinia e.g. tuberculosis particularly in  
 CC veterinary treatment. They are useful in screening assays for identifying  
 CC antimicrobial drugs. They are also used to prepare vaccines. The present  
 CC sequence is Yersinia pestis 2G5 mutant protein  
 XX  
 SQ Sequence 587 AA;  
 Query Match 78.0%; Score 39; DB 6; Length 587;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 1;  
 QY 3 WFEIDIV 9  
 DB 272 WFEIDIV 278

RESULT 6  
AAB20498  
ID AAB20498 standard; protein; 1717 AA.  
XX AC AAB20498;  
XX DT 06-AUG-2003 (revised)  
XX DT 09-JUL-2001 (first entry)  
XX DE Ostrinia nubilalis Bt toxin receptor.  
XX KW European corn borer; ECB; Bacillus thuringiensis; Bt toxin; receptor;  
XX KW crystal protein; CryIA; biological control; insecticide; crop protection.  
XX OS Ostrinia nubilalis.  
XX FH Key Location/Qualifiers  
FT Binding-site 1293..1462  
FT /label=CryIA\_binding\_site  
FT Domain 1571..1589  
FT /label=Transmembrane\_domain  
XX WO200136639-A2.  
XX PN 25-MAY-2001.  
XX PD 17-NOV-2000; 2000WO-US031674.  
XX PF 18-NOV-1999; 99US-0166285P.  
XX PR 21-SEP-2000; 2000US-0234099P.  
XX PP (PION-) PIONEER HI-BRED INT INC.  
XX PI Flannagan RD, Mathis JP, Meyer TE;  
XX WPI; 2001-329223/34.  
XX DR N-PSDB; AAF30933.  
XX PT New nucleic acid encoding Bt toxin receptor, useful for screening and  
PT identifying new Bt toxin receptor ligands useful as new insecticidal  
PT toxins.  
XX Claim 4(a); Page 57-61; 85pp; English.  
XX The present sequence of is that of the Bacillus thuringiensis (Bt) toxin  
XX receptor of the European corn borer (ECB), Ostrinia nubilalis. The  
XX sequence was deduced from cDNA isolated from an ECB larva midgut cDNA.  
XX The receptor protein binds CryIA toxins, particularly CryIA(b) toxins.  
XX The invention provides compositions and methods for modulating the  
XX susceptibility of a cell to Bt toxins. The compositions include Bt toxin  
XX receptor proteins from the lepidopteran insects ECB, corn earworm  
XX (Heliothis zea) and fall armyworm (Spodoptera frugiperda). Nucleic acids  
XX encoding the polypeptides, antibodies specific to the polypeptides, and  
XX nucleic acid constructs for expressing the polypeptides in cells are also  
XX provided. The methods of the invention are useful for investigating the  
XX structure-function relationships of Bt toxin receptors, investigating  
XX toxin-receptor interactions, elucidating the mode of action of Bt toxins,  
XX screening and identifying novel Bt toxin receptor ligands including novel  
XX insecticidal toxins, and designing and developing novel Bt toxin receptor  
XX ligands. The methods are useful for managing Bt toxin resistance in plant  
XX pests, and protecting plants against damage by plant pests. (Updated on  
XX 06-AUG-2003 to correct OS field.)  
SQ Sequence 1717 AA;  
Query Match 78.0%; Score 39; DB 4; Length 1717;  
Best Local Similarity 85.7%; Pred. No. 4.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLWFEID 7  
DB 650 YLWFEID 656

RESULT 7  
ABM69298  
ID ABM69298 standard; protein; 589 AA.  
XX AC ABM69298;  
XX DT 20-NOV-2003 (first entry)  
XX DE Photorhabdus luminescens protein sequence #2395.  
XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
XX KW detection; food; gene expression; plant; animal; microorganism; toxin;  
XX KW antibiotic; biopesticide; virulence factor; disease model; plague;  
XX KW whooping cough.  
XX OS Photorhabdus luminescens.  
XX PN WO200294867-A2.  
XX PD 28-NOV-2002.  
XX PF 07-FEB-2002; 2002WO-IB003040.  
XX PR 07-FEB-2001; 2001FR-00001659.  
XX PA (INSP ) INST PASTEUR.  
XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
XX PI Buchrieser C;  
XX WPI; 2003-148459/14.  
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
XX Claim 2; SEQ ID NO 2395; 1205pp; French.  
XX The invention relates to the isolation of genes and their encoded  
XX proteins from Photorhabdus luminescens. The isolated sequences are  
XX sources of probes and primers for detecting the genome of P. luminescens  
XX and related species; to study polymorphisms; for gene analysis and for  
XX detection/amplification of the genes. Antibodies (Ab) raised against the  
XX polypeptides encoded by the genes are used for detection/identification  
XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
XX carry a gene-containing vector are used to select compounds that  
XX modulate, regulate, induce or inhibit expression of the genes in plants,  
XX animals or microorganisms other than P. luminescens and are able to alter  
XX response or sensitivity to toxins and antibiotics produced by P.  
XX luminescens. Cells transformed to express the genes are useful for  
XX recombinant production of the proteins, particularly toxins and  
XX antibacterials useful as insecticides, bactericides and fungicides. The  
XX genes, proteins, vectors containing the genes and Ab are also useful  
XX therapeutically to treat microbial infection by bacteria or fungi that  
XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
XX biopesticides. Other uses of the genes and the proteins are as virulence  
XX factors and for identifying targets of human diseases for which P.  
XX luminescens is a model (particularly plague and whooping cough). This  
XX sequence represents one of the isolated P. luminescens proteins  
SQ Sequence 589 AA;  
Query Match 76.0%; Score 38; DB 6; Length 589;  
Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 WFEIDIV 9  
DB 273 WFEIDIV 279



## RESULT 8

ABB49338  
 ID ABB49338 standard; protein; 150 AA.  
 XX  
 AC ABB49338;  
 XX  
 DT 05-FEB-2002 (first entry)  
 XX  
 DE Listeria monocytogenes protein #2042.  
 XX  
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KW vitamin B12; bacterial infection; disease.  
 XX  
 OS Listeria monocytogenes.  
 XX  
 PN WO200177335-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 11-APR-2001; 2001WO-FR001118.  
 XX  
 PR 11-APR-2000; 2000FR-00004629.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;  
 PI Dussurget O, Chetoui F, Nedjari H, Glaeser P, Kunst F, Cossart P;  
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
 PI Rose M, Voss H;  
 XX  
 DR WPI; 2002-010914/01.  
 XX

Genomic sequence for Listeria monocytogenes, useful e.g. for treatment and prevention of Listeria and related bacterial infections, and related polypeptides.

Claim 6; SEQ ID NO 2043; 192pp; French.

The present invention relates to the genome sequence of Listeria monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of it are useful for selecting probes and primers for detecting genes in L. monocytogenes and related organisms, and for studying genetic polymorphisms and other genomes. The present sequence is a protein encoded by the genome sequence of the present invention. Proteins expressed from the genome sequence are useful for raising specific antibodies, identification of L. monocytogenes and related organisms, and for biosynthesis and biodegradation, especially biosynthesis of Vitamin B12. The genome sequence and proteins encoded by it are also useful for selecting compounds that regulate gene expression and cell replication and modulate L. monocytogenes-related diseases. In addition, the genome sequence and proteins encoded by it are useful in pharmaceutical and vaccines compositions for the treatment or prevention of infections by L. monocytogenes and related organisms. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 150 AA;

Query Match 74.0%; Score 37; DB 5; Length 150;

Best Local Similarity 77.8%; Pred. No. 80;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9

Db 116 FKWFEIDKV 124

## RESULT 9

Query Match 74.0%; Score 37; DB 5; Length 273;  
 Best Local Similarity 55.6%; Pred. No. 1.5e+02;

## ABP64842

ID ABP64842 standard; protein; 273 AA.  
 XX  
 AC ABP64842;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human protein SEQ ID 502.  
 XX  
 KW Human; expressed sequence tag; EST; haematopoietic disorder;  
 KW central nervous system disease; viral infection;  
 KW peripheral nervous system disease; non-healing wound; infectious disease;  
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
 KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
 KW anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;  
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;  
 KW immunostimulant; cerebroprotective.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200259260-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 16-NOV-2001; 2001WO-US042950.  
 XX  
 PR 17-NOV-2000; 2000US-00714936.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 XX  
 DR WPI; 2002-590824/63.  
 DR N-PSDB; ABQ99428.  
 XX  
 PT New isolated polynucleotide, useful in research, diagnostic or  
 PT therapeutic methods, e.g. preventing or treating disorders involving  
 PT aberrant protein expression or biological activity.  
 XX  
 PS Claim 20; SEQ ID NO 502; 394pp; English.  
 XX  
 CC The present invention relates to novel human coding sequences (ABQ99268-  
 CC ABQ99608) and proteins (ABP64842-ABP65022). The sequences are useful in  
 CC therapeutic, diagnostic and research methods. The polynucleotides may be  
 CC used in the field of molecular biology as hybridisation probes, primers  
 CC for PCR, for chromosome and gene mapping, for the recombinant production  
 CC of protein, or in generation of anti-sense DNA or RNA. The  
 CC polynucleotides are useful in diagnostics as expressed sequence tags  
 CC (ESTs) for identifying expressed genes or for physical mapping of the  
 CC human genome. The proteins may be used as molecular weight markers, or as  
 CC nutritional sources or supplements. The proteins may be used to maintain  
 CC and expand cell population in a totipotent or pluripotent state  
 CC useful for re-engineering damaged or diseased tissues, transplantation,  
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
 CC polynucleotides and proteins are useful for preventing, treating or  
 CC ameliorating disorders involving aberrant protein expression or  
 CC biological activity, e.g. haematopoietic disorders, central/peripheral  
 CC nervous system diseases, mechanical and traumatic disorders, non-healing  
 CC wounds, immune deficiencies and disorders, infectious diseases caused by  
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
 CC reactions and conditions, coagulation disorders, or cancer. The  
 CC polynucleotide sequences of the invention were assembled from ESTs  
 CC isolated mainly by sequencing by hybridisation, and in some cases,  
 CC sequences obtained from one or more public databases. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 273 AA;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 Db 10 FLWLQDLV 18

RESULT 10  
 ABB54865  
 ID ABB54865 standard; protein; 528 AA.  
 XX AC  
 XX ABBS4865;  
 DT 29-AUG-2003 (revised)  
 DT 16-MAY-2002 (first entry)  
 XX  
 DE Lactococcus lactis protein ypgD.  
 XX  
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.  
 XX  
 OS Lactococcus lactis; IL1403.  
 XX  
 PN FR2807446-A1.  
 XX PD  
 XX 12-OCT-2001.  
 XX PF  
 XX 11-APR-2000; 2000FR-00004630.  
 XX PR  
 XX 11-APR-2000; 2000FR-00004630.  
 XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;  
 XX WPI; 2002-043418/06.  
 XX  
 PT New nucleotide sequence useful in the identification of Lactococcus  
 PT lactis and related species.  
 XX  
 PS Claim 6; SEQ ID NO 1567; 2504pp; French.  
 XX  
 CC The present invention is related to a Lactococcus lactis nucleotide  
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic  
 CC acid sequence is useful in the detection and/or amplification of nucleic  
 CC acid sequence, particularly to identify Lactococcus lactis or related  
 CC species. The proteins of the invention are useful for the biosynthesis or  
 CC biodegradation of a composition of interest. The invention helps research  
 CC in lactic bacteria, particularly useful in the production of yogurt and  
 CC cheese. Note: The sequence data for this patent is based on equivalent  
 CC patent WO200177334 (published 18-OCT-2001) which is available in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 29-AUG-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 528 AA;

Query Match 74.0%; Score 37; DB 5; Length 528;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 Db 49 FLWLAIIDII 57

RESULT 11  
 ABM70700  
 ID ABM70700 standard; protein; 587 AA.  
 XX AC  
 XX ABM70700;  
 XX DT  
 DT 20-NOV-2003 (first entry)  
 XX

DE XX Photorhabdus luminescens protein sequence #3797.  
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 KW detection; food; gene expression; plant; animal; microorganism; toxin;  
 KW antibiotic; biopesticide; virulence factor; disease model; plague;  
 KW whooping cough.  
 XX  
 OS Photorhabdus luminescens.  
 XX  
 PN WO200294867-A2.  
 XX PD  
 XX 28-NOV-2002.  
 XX PF  
 PF 07-FEB-2002; 2002WO-IB003040.  
 XX PR  
 PR 07-FEB-2001; 2001FR-00001659.  
 XX PA (INSP ) INST PASTEUR.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunat F, Danchin A;  
 PI Buchrieser C;  
 XX WPI; 2003-148459/14.  
 DR  
 XX  
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
 XX  
 PS Claim 2; SEQ ID NO 3797; 1205pp; French.  
 XX  
 CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of P. luminescens  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than P. luminescens and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by P.  
 CC luminescens. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically to treat microbial infection by bacteria or fungi that  
 CC are sensitive to P. luminescens-encoded toxins or antibiotics and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which P.  
 CC luminescens is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated P. luminescens proteins  
 XX  
 SQ Sequence 587 AA;

Query Match 74.0%; Score 37; DB 6; Length 587;  
 Best Local Similarity 71.4%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEIDIV 9  
 Db 272 WFEIDIV 278

RESULT 12  
 ABB53431  
 ID ABB53431 standard; protein; 269 AA.  
 XX AC  
 XX ABB53431;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 16-MAY-2002 (first entry)  
 XX  
 DE Lactococcus lactis protein ybdC.



DR WPI: 2003-615309/58.  
 DR N-PSDB; ABD08859.  
 XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX  
 PS Disclosure; SEQ ID NO 24034; 455pp; English.  
 XX  
 CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
 CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-  
 CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX  
 SQ Sequence 373 AA;  
 Query Match 72.0%; Score 36; DB 7; Length 373;  
 Best Local Similarity 55.6%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLWFEIDIV 9  
 :|||:|:  
 Db 307 YLWVELDLV 315  
 RESULT 15  
 ABG01159  
 ID ABG01159 standard; protein; 410 AA.  
 XX  
 AC ABG01159;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #1150.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR N-PSDB; AAS65346.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX

PS Claim 20; SEQ ID NO 31518; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological actions. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 410 AA;  
 Query Match 72.0%; Score 36; DB 4; Length 410;  
 Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLWFEIDI 8  
 :|||:|:  
 Db 200 FLWFEEDDV 207  
 Search completed: December 30, 2004, 20:43:01  
 Job time : 13.4962 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:30:54 ; Search time 1.97229 Seconds  
(without alignments)  
439.058 Million cell updates/sec

Title: US-10-017-327-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	78.0	474	T38905	probable valine-py
2	39	78.0	587	A10180	conserved hypothet
3	37	74.0	110	S40136	T-cell receptor V-
4	37	74.0	134	C29774	T-cell receptor al
5	37	74.0	150	AG1585	hypothetical prote
6	37	74.0	150	AH1231	hypothetical prote
7	37	74.0	529	B86815	hypothetical prote
8	36	72.0	269	A86641	hypothetical prote
9	36	72.0	413	E84815	ybnc protein - Esc
10	36	72.0	413	D85587	probable synthetas
11	36	72.0	413	C90737	cardiolipin synthe
12	35	70.0	250	D69457	hypothetical prote
13	35	70.0	280	B82372	SpoOM-related prot
14	35	70.0	648	T08856	hypothetical prote
15	35	70.0	848	T25325	hypothetical prote
16	35	70.0	1190	T00842	probable histidine
17	35	70.0	1207	T52459	sensory transducti
18	34	68.0	121	A12244	dihydroneopterin a
19	34	68.0	175	A10185	probable membrane
20	34	68.0	209	F11612	glycosyl transfera
21	34	68.0	229	A81468	ribosomal protein
22	34	68.0	254	C90024	hypothetical prote
23	34	68.0	422	T05592	tyrosine transamin
24	34	68.0	469	S44620	C50C3.1 protein -
25	34	68.0	643	S55593	membrane protein S
26	34	68.0	912	T49399	hypothetical prote
27	34	68.0	2363	T38841	probable pre-mRNA
28	33	66.0	55	C82565	hypothetical prote
29	33	66.0	126	A82519	hypothetical prote

hypothetical prote  
dihydroneopterin a  
conserved hypothet  
hypothetical prote  
protein T10022.3 [  
hypothetical prote  
probable polyketid  
prolipoprotein dia  
delta-9 desaturase  
hypothetical prote  
hexosephosphate tr  
probable valine-py  
NADH2 dehydrogenas  
cellulose 1,4-beta  
hypothetical prote

ALIGNMENTS

RESULT 1

T38905  
probable valine-pyruvate transaminase (EC 2.6.1.66) SPAC56E4.03 [similarity] - fission y  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T38905  
R/Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A/Reference number: Z21813  
A/Accession: T38905  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-474 <CON>  
A/Cross-references: UNIPROT:O14192; EMBL:Z99261; NID:g4038621; PIDN:CAB16394.1; PID:g241  
A/Experimental source: strain 972h-; cosmid c56E4  
C/Genetics:  
A/Map position: 1  
A/Gene: SPDB:SPAC56E4.03  
C/Superfamily: Escherichia coli valine-pyruvate transaminase  
C/Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate  
F/297/Binding site: pyridoxal phosphate (lufs) (covalent) #status predicted

Query Match 78.0%; Score 39; DB 2; Length 474;  
Best Local Similarity 71.4%; Pred. No. 12;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7  
|:||||:  
Db 390 FIWFEVD 396

RESULT 2

A10180  
conserved hypothetical protein YP01485 [imported] - Yersinia pestis (strain CO92)  
C/Species: Yersinia pestis  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: A10180  
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A/Reference number: AB0001; MUID:21470413; PMID:11586360  
A/Accession: A10180  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-587 <KUR>  
A/Cross-references: UNIPROT:Q8ZG34; GB:AL590842; PIDN:CAC930308.1; PID:gl5979527; GSPDB:G  
C/Genetics:  
A/Gene: YP01485  
C/Superfamily: uncharacterized conserved protein

Query Match 78.0%; Score 39; DB 2; Length 587;  
 Best Local Similarity 85.7%; Pred. No. 15;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEIDIV 9  
 |||||:  
 Db 272 WFEIDV 278

RESULT 3  
 S40136  
 T-cell receptor V-alpha 8.1b - human  
 C:Species: Homo sapiens (man)  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
 C:Accession: S40136  
 R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.  
 Submitted to the EMBL Data Library, February 1993  
 A:Reference number: S40136  
 A:Accession: S40136  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-110 <PLA>  
 A:Cross-references: EMBL:X70308; NID:9437034; PIDN:CAA49787.1; PID:9437035  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: T-cell receptor

Query Match 74.0%; Score 37; DB 2; Length 110;  
 Best Local Similarity 55.6%; Pred. No. 5.8;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 ||| :|:  
 Db 10 FLWLQDLV 18

RESULT 4  
 C29774  
 T-cell receptor alpha chain precursor V region (HAP41) - human  
 C:Species: Homo sapiens (man)  
 C>Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-May-1997  
 C:Accession: C29774  
 R:Yoshikai, Y.; Kimura, N.; Toyonaga, B.; Mak, T.W.  
 J. Exp. Med. 164, 90-103, 1986  
 A:Title: Sequences and repertoire of human T cell receptor alpha chain variable region  
 A:Reference number: A92778; MUID:86253078; PMID:3088203  
 A:Accession: C29774  
 A:Molecule type: mRNA  
 A:Residues: 1-134 <YOS>  
 C:Genetics:  
 A:Map position: 14q11.2  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: T-cell receptor

Query Match 74.0%; Score 37; DB 2; Length 134;  
 Best Local Similarity 55.6%; Pred. No. 7.2;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 ||| :|:  
 Db 10 FLWLQDLV 18

RESULT 5  
 AG1585  
 hypothetical protein lin1224 [imported] - Listeria innocua (strain Clip11262)  
 C:Species: Listeria innocua  
 C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AG1585  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AG1585  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-150 <GLA>  
 A:Cross-references: UNIPROT:Q92CE8; GB:AL592022; PIDN:CAC96455.1; PID:gl6413698; GSPDB:G  
 A:Experimental source: strain Clip11262  
 C:Genetics:  
 A:Gene: lin1224

Query Match 74.0%; Score 37; DB 2; Length 150;  
 Best Local Similarity 77.8%; Pred. No. 8.1;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 |||||:  
 Db 116 FKWFEIDKV 124

RESULT 6  
 AH1231  
 hypothetical protein lmo1256 [imported] - Listeria monocytogenes (strain EGD-e)  
 C:Species: Listeria monocytogenes  
 C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AH1231  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.,  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AH1231  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-150 <GLA>  
 A:Cross-references: UNIPROT:Q8Y7L8; GB:NC\_003210; PIDN:CAC99334.1; PID:gl6410672; GSPDB:G  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo1256

Query Match 74.0%; Score 37; DB 2; Length 150;  
 Best Local Similarity 77.8%; Pred. No. 8.1;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 |||||:  
 Db 116 FKWFEIDKV 124

RESULT 7  
 B86815  
 hypothetical protein ypgp [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
 C:Species: Lactococcus lactis subsp. lactis  
 C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: B86815  
 R:Solotin, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich  
 Genome Res. 11, 731-753, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp  
 A:Reference number: A86625; MUID:21235186; PMID:11337471  
 A:Accession: B86815  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-529 <STO>  
 A:Cross-references: UNIPROT:Q9CF66; GB:AE005176; PID:gl2724521; PIDN:AAK05620.1; GSPDB:G  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: ypgp

Query Match 74.0%; Score 37; DB 2; Length 529;

```

Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9
Db 49 FLWLAIIDII 57

RESULT 8
A86641
hypothetical protein ybdc [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86641
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: A86641
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-269 <STO>
A;Cross-references: UNIPROT:Q9CJ72; GB:AE005176; PID:G12722977; PIDN:AAK04227.1; GSPDB:C
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ybdc
C;Superfamily: stage III sporulation protein; stage III sporulation protein homology

Query Match 72.0%; Score 36; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEI 6
Db 155 FLWFEI 160

RESULT 9
E64815
ybhO protein - Escherichia coli (strain K-12)
N;Alternate names: protein b0789
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: E64815
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E64815
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-413 <BLAT>
A;Cross-references: UNIPROT:P75771; GB:AE000181; GB:U00096; NID:G1786998; PIDN:AAC73876.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ybhO
C;Superfamily: cardiolipin synthase

Query Match 72.0%; Score 36; DB 2; Length 413;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8
Db 40 FIWFEDDV 47

RESULT 10
D85587
probable synthetase ybhO [imported] - Escherichia coli (strain O157:H7, substrain EDL933
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

```

```

C;Accession: D85587
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, B.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85587
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;Cross-references: UNIPROT:P75771; GB:AE005174; NID:G12513778; PIDN:AAG55160.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ybhO
C;Superfamily: cardiolipin synthase

Query Match 72.0%; Score 36; DB 2; Length 413;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8
Db 40 FIWFEDDV 47

RESULT 11
C90737
cardiolipin synthetase [imported] - Escherichia coli (strain O157:H7, substrain RMD 050
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C90737
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C90737
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <HAY>
A;Cross-references: UNIPROT:P75771; GB:BA000007; PIDN:BA34290.1; PID:G13360326; GSPDB:G
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs0867
C;Superfamily: cardiolipin synthase

Query Match 72.0%; Score 36; DB 2; Length 413;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8
Db 40 FIWFEDDV 47

RESULT 12
D69457
hypothetical protein AF1661 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: D69457
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleisichmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: D69457
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-250 <KLE>

```

A;Cross-references: UNIPROT:O28612; GB:AE000989; GB:AE000782; NID:g2689312; PIDN:AAB8961

Query Match 70.0%; Score 35; DB 2; Length 250;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LWFPEIDI 8  
 Db 101 YLWFAVDV 108  
 :|||:|:

## RESULT 13

B82372  
 SpOw-related protein VC0039 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C;Species: Vibrio cholerae  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C;Accession: B82372  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82372

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-280 <HEI>

A;Cross-references: UNIPROT:Q9KVV0; GB:AE004096; GB:AE003852; NID:g9654440; PIDN:AAF9321  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0039

A;Map position: 1

C;Superfamily: Bacillus subtilis hypothetical protein ygai

Query Match 70.0%; Score 35; DB 2; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 38;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFPEID 7  
 Db 229 LWFPEVD 234  
 |||||:

## RESULT 14

T08856  
 hypothetical protein A\_TM017A05.5 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C;Accession: T08856  
 R;Waterston, R.

submitted to the EMBL Data Library, October 1997

A;Reference number: Z16500

A;Accession: T08856

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-648 <WAT>

A;Cross-references: UNIPROT:Q95XL4; EMBL:AF024504; NID:g2435510; PID:g2435516

A;Experimental source: cultivar Columbia

C;Genetics:

A;Map position: 4

A;Introns: 47/3; 90/1; 120/1; 185/1; 208/2; 222/3; 570/3

A;Note: A\_TM017A05.5

Query Match 70.0%; Score 35; DB 2; Length 648;  
 Best Local Similarity 83.3%; Pred. No. 93;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFPEID 7  
 Db 175 LWFPEVD 180  
 |||||:

## RESULT 15

T25325

hypothetical protein T26H2.7 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T25325

R;Matthews, L.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z20016

A;Accession: T25325

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-848 <WIL>

A;Cross-references: UNIPROT:O18139; EMBL:Z82055; PIDN:CAB04848.1; GSPDB:GN00023; CESP:T26H2

A;Experimental source: clone T26H2

C;Genetics:

A;Gene: CESP:T26H2.7

A;Map position: 5

A;Introns: 41/2; 129/3; 176/1; 254/2; 312/1; 388/1; 425/3; 466/1; 514/3; 557/3; 628/3; 61

C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 70.0%; Score 35; DB 2; Length 848;

Best Local Similarity 71.4%; Pred. No. 1.2e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFPEIDI 8

Db 699 LWFQLDI 705

|||||:

Search completed: December 30, 2004, 20:52:21

Job time : 2.97229 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:29:59 ; Search time 10.5642 Seconds  
(without alignments)  
490.180 Million cell updates/sec

Title: US-10-017-327-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	78.0	266	2 Q6D9R7	Q6D9R7 erwinia car
2	39	78.0	474	2 O14192	O14192 schizosacch
3	39	78.0	587	2 Q8ZG34	Q8ZG34 yersinia pe
4	39	78.0	587	2 AAS61616	AAS61616 yersinia
5	39	78.0	596	2 Q8IIA6	Q8IIA6 plasmodium
6	39	78.0	1714	2 Q6IZ03	Q6IZ03 ostrinia nu
7	39	78.0	1714	2 AAT37678	AAT37678 ostrinia
8	38	76.0	588	2 Q7MZS2	Q7MZS2 photorhabd
9	37	74.0	150	2 Q92CE8	Q92CE8 listeria in
10	37	74.0	150	2 Q8Y7L8	Q8Y7L8 listeria mo
11	37	74.0	198	2 Q7RSS9	Q7RSS9 plasmodium
12	37	74.0	322	2 Q7UJ06	Q7UJ06 rhodospirell
13	37	74.0	529	2 Q9CFF6	Q9CFF6 lactococcus
14	37	74.0	586	2 Q7N267	Q7N267 photorhabd
15	37	74.0	713	2 Q7RVSA	Q7RVSA neurospora
16	37	74.0	731	2 Q74693	Q74693 neurospora
17	36	72.0	129	2 Q8DKV4	Q8DKV4 synecococc
18	36	72.0	264	2 Q8BDE9	Q8BDE9 vibrio vuln
19	36	72.0	266	2 Q7MGK0	Q7MGK0 vibrio vuln
20	36	72.0	269	1 OXAI LACLA	OXAI LACLA
21	36	72.0	270	2 Q87KC8	Q87KC8 lactococcus
22	36	72.0	331	2 Q93X27	Q93X27 populus eur
23	36	72.0	413	1 YBHO_ECOL6	YBHO_ECOL6
24	36	72.0	413	1 YBHO_ECOL1	YBHO_ECOL1
25	36	72.0	445	2 Q70XS6	Q70XS6 escherichia
26	36	72.0	445	2 CAD45022	CAD45022 scutigera
27	36	72.0	590	1 SYD_PHOLL	SYD_PHOLL
28	36	72.0	700	1 HPSI_HUMAN	HPSI_HUMAN
29	36	72.0	700	2 Q8WXE5	Q8WXE5 homo sapien
30	36	72.0	704	1 HPSI_MOUSE	HPSI_MOUSE
31	36	72.0	706	2 Q99MK7	Q99MK7 rattus norv

32	35	70.0	76	2 Q935T6	Q935T6 streptococc
33	35	70.0	112	2 Q9WU87	Q9WU87 mus musculu
34	35	70.0	194	2 Q8VEN8	Q8VEN8 mus musculu
35	35	70.0	205	2 Q83AZ7	Q83AZ7 coxiella bu
36	35	70.0	232	2 Q8IXE0	Q8IXE0 homo sapien
37	35	70.0	250	1 Y661_ARCFU	Y661_ARCFU
38	35	70.0	261	2 Q8ES40	Q8ES40 oceanobacil
39	35	70.0	280	2 Q9KVV0	Q9KVV0 vibrio chol
40	35	70.0	285	1 FFG_COREF	FFG_COREF
41	35	70.0	309	2 Q7TRM3	Q7TRM3 corynebacte
42	35	70.0	309	2 Q8VFT9	Q8VFT9 mus musculu
43	35	70.0	311	2 Q7TRM1	Q7TRM1 mus musculu
44	35	70.0	311	2 Q7TRM2	Q7TRM2 mus musculu
45	35	70.0	311	2 Q8VF70	Q8VF70 mus musculu

#### ALIGNMENTS

##### RESULT 1

Q6D9R7 ID Q6D9R7 PRELIMINARY; PRT; 266 AA.  
AC Q6D9R7;  
DT 01-OCT-2004 (Tremblrel. 28, Created)  
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)  
DE 01-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Putative plasmid transfer protein.  
GN Name=traE; ORFNames=ECA0548;  
OS Erwinia carotovora subsp. atroseptica SCRI1043.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=218491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI1043;  
RA Bell K.S., Sebahia M., Pritchard L., Holden M., Hyman L.J.,  
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,  
RA Atkin R., Bacon N., Brooks K., Chillingworth T., Clark K., Doggett J.,  
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,  
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,  
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
SQ EMBL; BX950851; CAG73463.1; -  
SQ SEQUENCE 266 AA; 30560 MW; 8D7628F313223FBE CRC64;

Query Match 78.0%; Score 39; DB 2; Length 266;  
Best Local Similarity 85.7%; Pred. No. 41;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFEIDI 8  
|||  
Db 134 LWFEIDV 140

##### RESULT 2

O14192 ID O14192 PRELIMINARY; PRT; 474 AA.  
AC O14192;  
DT 01-JAN-1999 (Tremblrel. 09, Created)  
DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE SPAC56E4.03 protein (EC 2.6.1.-).  
GN Name=SPAC56E4.03;  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetaceae;  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gencies S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
RA Weidjens I., Vanstaels E., Rieger M., Schafer M., Muller-Auer S.,  
RA Gabel C., Fuchs M., Dusterhoft A., Fritzc C., Holzer E., Moestl D.,  
RA Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,  
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin J., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
DR EMBL; Z99261; CAB16394.1; --  
DR PIR; T38905; T38905.  
DR GenDB SPombs; SPAC56E4.03; --  
DR GO; GO:0008483; F:transaminase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR002052; N6 Mtase.  
DR PROSITE; PS00092; N6 MTASE; UNKNOWN\_1.  
KW Amino transferase; Transferase.  
SQ SEQUENCE 474 AA; 53166 MW; E546EABC7E884623 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 474;  
Best Local Similarity 71.4%; Pred. No. 74;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7  
Db 390 FIMFEVD 396  
|:||||:|  
|:||||:|

RESULT 3  
Q8ZG34 PRELIMINARY; PRT; 587 AA.  
AC Q8ZG34; Q74VD6; Q7CHP0;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein YP01485.  
GN OrderedLocusNames=Yp1375, YP01485, Y2684;  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebaiha M., James K.D., Churcher C.M., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.-M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Hollroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIMS / Biovar Mediaevalis;

RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM.";  
RL J. Bacteriol. 184:4601-4611(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91001 / Biovar Mediaevalis;  
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,  
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,  
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
RA Yang R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ414148; CAC90308.1; --  
DR EMBL; AE013870; AAM86237.1; --  
DR EMBL; AE017132; AAS61616.1; --  
DR PIR; A10180; A10180. DUF879.  
DR InterPro; IPR010272; DUF879.  
DR Pfam; PF05947; DUF879; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 587 AA; 67418 MW; 389A200B93753764 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 587;  
Best Local Similarity 85.7%; Pred. No. 91;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
Db 272 WFEIDVV 278  
|||||:  
|||||:

RESULT 4  
AAS61616 PRELIMINARY; PRT; 587 AA.  
AC AAS61616;  
DT 24-MAR-2004 (TrEMBLrel. 27, Created)  
DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN YP1375.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91001 / Biovar Mediaevalis;  
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,  
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,  
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
RA Yang R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE017132; AAS61616.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 587 AA; 67418 MW; 389A200B93753764 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 587;  
Best Local Similarity 85.7%; Pred. No. 91;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
Db 272 WFEIDVV 278  
|||||:  
|||||:

RESULT 5  
Q8IIA6 PRELIMINARY; PRT; 596 AA.  
AC Q8IIA6;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=PF11\_0268;  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22255705; PubMed=12368864;  
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
 RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,  
 RA Perlee M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,  
 RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
 RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
 RA Fraser C.M., Barrell B.;  
 RT "Genome sequence of the human malaria parasite Plasmodium  
 falciparum.";  
 RL Nature 419:498-511(2002).  
 DR EMBL; AE014839; AAN35852.1; -;  
 DR InterPro; IPR011043; Gal\_oxid\_central.  
 DR InterPro; IPR011498; Kelch\_2.  
 DR Pfam; PF01344; Kelch\_1; 2.  
 DR Pfam; PF07646; Kelch\_2; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 596 AA; 68046 MW; B07DCS8441A754AF CRC64;

Query Match 78.0%; Score 39; DB 2; Length 596;  
 Best Local Similarity 75.0%; Pred. No. 92;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8  
 Db 433 FSWFEVDI 440

RESULT 6  
 Q61Z03 PRELIMINARY; PRT; 1714 AA.  
 AC Q61Z03;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Cadherin Al.  
 OS Ostrinia nubilalis (European corn borer).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea;  
 OC Pyralidae; Pyraustinae; Ostrinia.  
 OX NCBI\_TaxID=29057;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Coates B.S., Sumnerford D.V., Hellmich R.L.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 11 cadherin domains.  
 DR EMBL; AY612336; AAT37678.1; -;  
 DR InterPro; IPR002126; Cadherin.  
 DR Pfam; PF00028; Cadherin; 5.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 10.  
 DR PROSITE; PS00232; CADHERIN\_1; 2.  
 DR PROSITE; PS00268; CADHERIN\_2; 11.  
 KW Calcium; Calcium-binding.  
 SQ SEQUENCE 1714 AA; 191900 MW; 89E3CAE0614654DE CRC64;

Query Match 78.0%; Score 39; DB 2; Length 1714;  
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7

Db 650 YLWFEID 656  
 :|||||

RESULT 7  
 AAT37678 PRELIMINARY; PRT; 1714 AA.  
 ID AAT37678  
 AC AAT37678;  
 DT 01-JUN-2004 (TReMBLrel. 27, Created)  
 DT 01-JUN-2004 (TReMBLrel. 27, Last sequence update)  
 DT 01-JUN-2004 (TReMBLrel. 27, Last annotation update)  
 DE Cadherin Al.  
 OS Ostrinia nubilalis (European corn borer).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea;  
 OC Pyralidae; Pyraustinae; Ostrinia.  
 OX NCBI\_TaxID=29057;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Coates B.S., Sumnerford D.V., Hellmich R.L.;  
 RL "Single nucleotide polymorphism (SNP) detection in proximity of two  
 putative toxin-binding regions of the candidate *Bacillus thuringiensis*  
 resistance gene, cadherin, of *Ostrinia nubilalis*.";  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY612336; AAT37678.1; -;  
 SQ SEQUENCE 1714 AA; 191900 MW; 89E3CAE0614654DE CRC64;

Query Match 78.0%; Score 39; DB 2; Length 1714;  
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7  
 Db 650 YLWFEID 656  
 :|||||

RESULT 8  
 Q7MZS2 PRELIMINARY; PRT; 588 AA.  
 ID Q7MZS2  
 AC Q7MZS2;  
 DT 01-MAR-2004 (TReMBLrel. 26, Created)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Similar to unknown protein.  
 GN OrderedLocNames=plu4202;  
 OS Photobacterium luminescens (subsp. laumondii).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photobacterium.  
 OX NCBI\_TaxID=141679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=TT01;  
 RX MEDLINE=22957627; PubMed=14528314;  
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,  
 RA Taurut S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,  
 RA Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,  
 RA Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,  
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;  
 RT "The genome sequence of the entomopathogenic bacterium *Photobacterium*  
*luminescens*.";  
 RL Nat. Biotechnol. 21:1307-1313(2003).  
 DR EMBL; BX571873; CAE16574.1; -;  
 DR Photolista; plu4202; -;  
 DR InterPro; IPR010272; DUF879.  
 DR Pfam; PF05947; DUF879; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 588 AA; 67517 MW; F5D33F2A6A5DFD5A CRC64;

Query Match 76.0%; Score 38; DB 2; Length 588;  
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9

```

Db      273 WFLDIV 279
||||:||||
ID Q92CE8      PRELIMINARY;      PRT;      150 AA.
AC Q92CE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lin1224 protein.
GN OrderedLocusNames=lin1224;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-Lopez N., Hain T., Hauf J., Jackson D.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- SIMILARITY: Belongs to the Nudix hydrolase family.
DR EMBL; AL596168; CAC96455.1; -.
DR FIRM; AG1585; AG1585.
DR ListList; LIN1224; -.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR Pfam; PF00293; NUDIX; 1.
DR PROSITE; PS00893; NUDIX; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 150 AA; 17636 MW; 205F376076CB8C87 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 150;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 FLWFEIDIV 9
Db      116 FKWFEIDKV 124
||||:|||||
ID Q7RSS9      PRELIMINARY;      PRT;      198 AA.
AC Q7RSS9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY00274;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perte M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; ABL01000077; EAA22282.1; -.
KW Hypothetical protein.
SQ SEQUENCE 198 AA; 22441 MW; 50A2D7E56AB5E236 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 198;
Best Local Similarity 75.0%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FLWFEIDI 8
Db      33 FSWYEIDI 40
||||:|||||

```

RESULT 12

```

Q7UJ06
ID Q7UJ06 PRELIMINARY; PRT; 322 AA.
AC Q7UJ06;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=R812213;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner P.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294154; CAD77454.1;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 322 AA; 35597 MW; 9790D69F2C0EB08B CRC64;

Query Match 74.0%; Score 37; DB 2; Length 322;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7
Db 79 FIWFID 85

RESULT 13
Q9CFP6
ID Q9CFP6 PRELIMINARY; PRT; 529 AA.
AC Q9CFP6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter ATP binding and permease protein.
GN Name=ypgD; OrderedLocusNames=LI1522;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE006383; AAK05620.1; -.
DR PIR; B66815; B66815.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane.1.
DR InterPro; IPR001140; ABC TM transp.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00664; ABC membrane.1.
DR Pfam; PF00005; ABC_tran.1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS50929; ABC_TM1F; 1.

DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 529 AA; 59763 MW; 7838906A4D09E0C5 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 529;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEID 9
Db 49 FLWLAIID 57

RESULT 14
Q7N267
ID Q7N267 PRELIMINARY; PRT; 586 AA.
AC Q7N267;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to unknown protein.
GN OrderedLocusNames=plu3225;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571869; CAE15599.1; -.
DR Photolact; plu3225; -.
DR InterPro; IPR010272; DUF879.
DR Pfam; PF05947; DUF879; 1.
KW Complete proteome.
SQ SEQUENCE 586 AA; 67513 MW; AA8AB383C031F30F CRC64;

Query Match 74.0%; Score 37; DB 2; Length 586;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9
Db 272 WFELDVV 278

RESULT 15
Q7RV54
ID Q7RV54 PRELIMINARY; PRT; 713 AA.
AC Q7RV54;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (AF088906) clock-controlled gene-9
DE protein.
GN Name=NCU09559.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rahman B.,

```

RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kamysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,  
 RA Kryetofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,  
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Varden O., Planann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,  
 RT "The Genome Sequence of the Filamentous Fungus *Neurospora crassa*."  
 RL Nature 0:0-0(2003).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABX01000379; EAA30248.1; -.  
 DR GO; GO:0009059; P:biosynthesis; IEA.  
 DR InterPro; IPR001296; Glyco trans 1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 713 AA; 79717 MW; DA651BA0E524DBEC CRC64;

Query Match 74.0%; Score 37; DB 2; Length 713;  
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LWFEIDIV 9  
 Db 124 LWLEVDIV 131

Search completed: December 30, 2004, 20:50:52  
 Job time : 12.5642 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:34:10 ; Search time 2.06297 seconds  
(without alignments)  
289.321 Million cell updates/sec

Title: US-10-017-327-9  
Perfect score: 45  
Sequence: 1 FLSYDLFVV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	77.8	65	US-08-905-223-453	Sequence 453, App
2	35	77.8	85	US-09-513-999C-5304	Sequence 5304, App
3	35	77.8	162	US-09-663-600A-126	Sequence 126, App
4	35	77.8	162	US-09-663-600A-220	Sequence 220, App
5	35	77.8	163	US-09-149-476-462	Sequence 462, App
6	35	77.8	203	US-09-766-055A-4	Sequence 4, Appli
7	35	77.8	231	US-09-270-767-38981	Sequence 38981, A
8	35	77.8	231	US-09-270-767-54198	Sequence 54198, A
9	35	77.8	382	US-09-248-796A-18423	Sequence 18423, A
10	34	75.6	108	US-09-270-767-45072	Sequence 45072, A
11	34	75.6	195	US-09-270-767-35680	Sequence 35680, A
12	34	75.6	195	US-09-270-767-50897	Sequence 50897, A
13	34	75.6	363	US-09-328-352-5693	Sequence 5693, Ap
14	33	73.3	103	US-09-543-681A-4764	Sequence 4764, Ap
15	33	73.3	370	US-08-837-593-7	Sequence 7, Appli
16	33	73.3	593	US-09-248-796A-19856	Sequence 19856, A
17	31	68.9	132	US-09-270-767-36909	Sequence 36909, A
18	31	68.9	132	US-09-270-767-52126	Sequence 52126, A
19	31	68.9	158	US-09-583-110-3871	Sequence 3871, Ap
20	31	68.9	195	US-09-270-767-38485	Sequence 38485, A
21	31	68.9	195	US-09-270-767-53702	Sequence 53702, A
22	31	68.9	246	US-09-583-110-4242	Sequence 4242, Ap
23	31	68.9	335	US-09-248-796A-17099	Sequence 17099, A
24	31	68.9	344	US-09-107-532A-5546	Sequence 5546, Ap
25	31	68.9	346	US-09-248-796A-15456	Sequence 15456, A
26	31	68.9	349	US-09-328-352-7078	Sequence 7078, Ap
27	31	68.9	400	US-09-107-532A-6266	Sequence 6266, Ap

28	31	68.9	437	4	US-09-540-236-2207	Sequence 2207, Ap
29	31	68.9	451	3	US-09-184-964-3	Sequence 3, Appli
30	31	68.9	453	4	US-09-538-092-472	Sequence 472, App
31	31	68.9	472	4	US-09-486-192-3	Sequence 3, Appli
32	31	68.9	481	3	US-09-537-357-7	Sequence 7, Appli
33	31	68.9	487	3	US-09-537-357-15	Sequence 15, Appli
34	30	66.7	75	4	US-09-134-000C-5626	Sequence 5626, Ap
35	30	66.7	123	4	US-09-248-796A-24081	Sequence 24081, A
36	30	66.7	126	4	US-09-889-463A-6	Sequence 6, Appli
37	30	66.7	130	4	US-09-198-452A-528	Sequence 528, App
38	30	66.7	156	4	US-09-270-767-43152	Sequence 43152, A
39	30	66.7	215	4	US-09-248-796A-28019	Sequence 28019, A
40	30	66.7	324	4	US-09-198-452A-724	Sequence 724, App
41	30	66.7	333	3	US-09-230-637-30	Sequence 30, Appli
42	30	66.7	343	4	US-09-489-039A-10892	Sequence 10892, A
43	30	66.7	346	4	US-09-543-681A-6716	Sequence 6716, Ap
44	30	66.7	349	1	US-08-343-027A-12	Sequence 12, Appli
45	30	66.7	402	4	US-09-107-532A-5454	Sequence 5454, Ap

ALIGNMENTS

RESULT 1  
US-08-905-223-453  
; Sequence 453, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 453:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 65 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: -21...-1  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 5.6  
; OTHER INFORMATION: seq LVLSQLFLLSYD/LF  
US-08-905-223-453

Query Match 77.8%; Score 35; DB 3; Length 65;  
Best Local Similarity 100.0%; Pred. No. 9.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
| | | | |  
Db 18 LSYDLFV 24

RESULT 2  
US-09-513-999C-5304  
; Sequence 5304, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5304  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 68  
; OTHER INFORMATION: Xaa=Ala or Pro or Ser or Thr

US-09-513-999C-5304

Query Match 77.8%; Score 35; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
| | | | |  
Db 18 LSYDLFV 24

RESULT 3  
US-09-663-600A-126  
; Sequence 126, Application US/09663600A  
; Patent No. 6573068  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bouquelert, Lydie  
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
; FILE REFERENCE: 31.US3.CIP  
; CURRENT APPLICATION NUMBER: US/09/663,600A  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/099,273  
; NUMBER OF SEQ ID NOS: 229  
; SOFTWARE: Patent.pm

US-09-663-600A-126

; SEQ ID NO 126  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -21...-1  
US-09-663-600A-126

Query Match 77.8%; Score 35; DB 4; Length 162;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
| | | | |  
Db 18 LSYDLFV 24

RESULT 4  
US-09-663-600A-220  
; Sequence 220, Application US/09663600A  
; Patent No. 6573068  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bouquelert, Lydie  
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
; FILE REFERENCE: 31.US3.CIP  
; CURRENT APPLICATION NUMBER: US/09/663,600A  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/099,273  
; NUMBER OF SEQ ID NOS: 229  
; SOFTWARE: Patent.pm

US-09-663-600A-220

Query Match 77.8%; Score 35; DB 4; Length 162;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
| | | | |  
Db 18 LSYDLFV 24

RESULT 5  
US-09-149-476-462  
; Sequence 462, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: PZ002PI  
; CURRENT APPLICATION NUMBER: US/09/149,476

US-09-149-476-462



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89</											

[illegible]

; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,589  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,593  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,614  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,578  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,576  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/047,501  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,670  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/056,632  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,664  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,876  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,881  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,909  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,875  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,862  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,887  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,908  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/057,650  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/056,884  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,669  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/049,610  
; EARLIER FILING DATE: 1997-06-13  
; EARLIER APPLICATION NUMBER: 60/061,060  
; EARLIER FILING DATE: 1997-10-02

Query Match 77.8%; Score 35; DB 4; Length 163;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

## RESULT 6

US-09-766-055A-4  
; Sequence 4, Application US/09766055A  
; Patent No. 6403354  
; GENERAL INFORMATION:  
; APPLICANT: XU, SHUANG-YONG  
; APPLICANT: SAMUELSON, JAMES  
; APPLICANT: PELLETIER, JOHN  
; APPLICANT: SIBLEY, MARION  
; APPLICANT: WILSON, GEOFFREY G.  
; TITLE OF INVENTION: METHOD FOR CLONING AND EXPRESSION OF BstYI RESTRICTION  
; TITLE OF INVENTION: ENDONUCLEASE AND BstYI METHYLASE IN E. coli AND  
; FILE REFERENCE: PURIFICATION OF BstYI AND M.BstYI ENZYMES  
; CURRENT APPLICATION NUMBER: US/09/766,055A  
; CURRENT FILING DATE: 2001-01-19

; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Bacillus stearothermophilus Y406  
US-09-766-055A-4

Query Match 77.8%; Score 35; DB 4; Length 203;  
Best Local Similarity 75.0%; Pred. No. 29;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
Db 136 FVAYDLFV 143

## RESULT 7

US-09-270-767-38981  
; Sequence 38981, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38981  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-38981

Query Match 77.8%; Score 35; DB 4; Length 231;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 9  
Db 172 ISYDLFII 179

## RESULT 8

US-09-270-767-54198  
; Sequence 54198, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 54198  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-54198

Query Match 77.8%; Score 35; DB 4; Length 231;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 9  
Db 172 ISYDLFII 179

```
RESULT 9
US-09-248-796A-18423
; Sequence 18423, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18423
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18423
Query Match 77.8%; Score 35; DB 4; Length 382;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9
Db 230 FLSYDLFV 238

RESULT 10
US-09-270-767-45072
; Sequence 45072, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45072
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45072
Query Match 75.6%; Score 34; DB 4; Length 108;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9
Db 86 FFSYDIFV 94

RESULT 11
US-09-270-767-35680
; Sequence 35680, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35680
```

```
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35680
Query Match 75.6%; Score 34; DB 4; Length 195;
Best Local Similarity 44.4%; Pred. No. 43;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9
Db 66 YISYELFV 74

RESULT 12
US-09-270-767-50897
; Sequence 50897, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50897
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50897
Query Match 75.6%; Score 34; DB 4; Length 195;
Best Local Similarity 44.4%; Pred. No. 43;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9
Db 66 YISYELFV 74

RESULT 13
US-09-328-352-5693
; Sequence 5693, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GFC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5693
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5693
Query Match 75.6%; Score 34; DB 4; Length 363;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9
Db 39 FLVYDIFV 47

RESULT 14
```

```
US-09-543-681A-4764
; Sequence 4764, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4764
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4764

Query Match      73.3%; Score 33; DB 4; Length 103;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LSYDLFV 8
Db      39 ISYDLFV 45

RESULT 15
US-08-837-593-7
; Sequence 7, Application US/08837593
; Patent No. 5977442
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Zhang Zhuqun
; TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced
; TITLE OF INVENTION: Map Kinase and its Use for Enhanced Disease Resistance in Plant
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
; ADDRESSEE: P.C.
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,593
; FILING DATE: April 21, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029,805
; FILING DATE: October 25, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; REFERENCE/DOCKET NUMBER: Rutgers 97-0016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-837-593-7
```

```
Query Match      73.3%; Score 33; DB 2; Length 370;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FLSYDLF 7
Db      21 FISYDIF 27
```

```
Search completed: December 30, 2004, 20:53:57
Job time : 3.06297 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:35:11 ; Search time 8.81864 Seconds  
(without alignments)  
367.126 Million cell updates/sec

Title: US-10-017-327-9  
Perfect score: 45  
Sequence: 1 FLSYDLFW 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues  
Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pcp.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	11	US-09-870-216C-9
2	45	100.0	9	13	US-10-017-327-9
3	36	80.0	51	11	US-09-833-245-2059
4	36	80.0	51	14	US-10-060-255-62
5	36	80.0	51	17	US-10-881-088-62
6	35	77.8	130	9	US-09-729-674-16
7	35	77.8	162	11	US-09-978-360A-732
8	35	77.8	162	14	US-10-319-763-126
9	35	77.8	162	14	US-10-319-763-220
10	35	77.8	162	14	US-10-170-385-53
11	35	77.8	162	15	US-10-264-237-1995
12	35	77.8	163	10	US-09-809-391-462
13	35	77.8	163	10	US-09-882-171-462

14	35	77.8	163	14	US-10-164-861-462	Sequence 462, App
15	34	75.6	55	9	US-09-867-550-916	Sequence 916, App
16	34	75.6	66	15	US-10-424-599-175697	Sequence 175697,
17	34	75.6	98	14	US-10-106-698-4683	Sequence 4683, Ap
18	34	75.6	247	15	US-10-108-260A-3015	Sequence 3015, Ap
19	34	75.6	346	15	US-10-424-599-217530	Sequence 217530,
20	34	75.6	412	10	US-09-882-227-396	Sequence 396, App
21	34	75.6	674	14	US-10-108-260A-3829	Sequence 3829, Ap
22	34	75.6	674	14	US-10-094-749-2686	Sequence 2686, Ap
23	33	73.3	56	15	US-10-424-599-278654	Sequence 278654,
24	33	73.3	67	15	US-10-424-599-284635	Sequence 284635,
25	33	73.3	88	17	US-10-425-115-318398	Sequence 318398,
26	33	73.3	100	17	US-10-425-115-195354	Sequence 195354,
27	33	73.3	211	14	US-10-083-357-691	Sequence 691, App
28	33	73.3	573	14	US-10-369-493-2314	Sequence 2314, Ap
29	33	73.3	573	14	US-10-369-493-2320	Sequence 2320, Ap
30	33	73.3	583	14	US-10-369-493-22678	Sequence 22678, A
31	33	73.3	839	15	US-10-424-599-206656	Sequence 206656,
32	33	73.3	1507	15	US-10-424-599-206659	Sequence 206659,
33	32	71.1	62	16	US-10-437-963-177564	Sequence 177564,
34	32	71.1	80	17	US-10-425-115-256908	Sequence 256908,
35	32	71.1	153	15	US-10-424-599-147896	Sequence 147896,
36	32	71.1	165	14	US-10-156-761-8778	Sequence 8778, Ap
37	32	71.1	265	15	US-10-336-603A-52	Sequence 52, Appl
38	32	71.1	279	9	US-09-778-927A-64	Sequence 64, Appl
39	32	71.1	288	15	US-10-336-603A-48	Sequence 48, Appl
40	32	71.1	291	16	US-10-437-963-131686	Sequence 131686,
41	32	71.1	305	15	US-10-336-603A-46	Sequence 46, Appl
42	32	71.1	313	17	US-10-739-930-6831	Sequence 6831, Ap
43	32	71.1	328	15	US-10-336-603A-44	Sequence 44, Appl
44	32	71.1	328	15	US-10-336-603A-50	Sequence 50, Appl
45	32	71.1	337	16	US-10-437-963-113203	Sequence 113203,

ALIGNMENTS

RESULT 1  
US-09-870-216C-9  
; Sequence 9, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-216C-9

Query Match 100.0%; Score 45; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLSYDLFW 9  
| | | | |  
Db 1 FLSYDLFW 9

RESULT 2  
US-10-017-327-9  
; Sequence 9, Application US/10017327



; CURRENT FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: 09/539,330  
; PRIOR FILING DATE: 2000-03-30  
; NUMBER OF SEQ ID NOS: 283  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-729-674-16

Query Match 77.8%; Score 35; DB 9; Length 130;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
|||  
Db 18 LSYDLFV 24

## RESULT 7

US-09-978-360A-732  
; Sequence 732, Application US/09978360A  
; Publication No. US20040110939A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; APPLICANT: Clusel, Catherine  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides

; FILE REFERENCE: 56.US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273  
; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PC7/IB98/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155  
; PRIOR FILING DATE: 1999-02-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810  
; SOFTWARE: Patent.pm  
; SEQ ID NO 732  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -94...-1  
US-09-978-360A-732

Query Match 77.8%; Score 35; DB 11; Length 162;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
|||  
Db 18 LSYDLFV 24

## RESULT 8

US-10-319-763-126  
; Sequence 126, Application US/10319763  
; Publication No. US20030144490A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS

; FILE REFERENCE: G-031.US04.DIV  
; CURRENT APPLICATION NUMBER: US/10/319,763  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/099,273  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 229  
; SOFTWARE: Patent.pm  
; SEQ ID NO 126  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -21...-1  
US-10-319-763-126

Query Match 77.8%; Score 35; DB 14; Length 162;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
|||  
Db 18 LSYDLFV 24

## RESULT 9

US-10-319-763-220  
; Sequence 220, Application US/10319763  
; Publication No. US20030144490A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
; FILE REFERENCE: G-031.US04.DIV  
; CURRENT APPLICATION NUMBER: US/10/319,763  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/099,273  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 229  
; SOFTWARE: Patent.pm  
; SEQ ID NO 220

```
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -94...-1
US-10-319-763-220

Query Match          77.8%; Score 35; DB 14; Length 162;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8
Db 18 LSYDLFV 24

RESULT 10
US-10-170-385-53
; Sequence 53, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; CURRENT APPLICATION NUMBER: US/10/170,385
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-53

Query Match          77.8%; Score 35; DB 14; Length 162;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8
Db 18 LSYDLFV 24

RESULT 11
US-10-264-237-1995
; Sequence 1995, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1

; SEQ ID NO 1995
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-1995

Query Match          77.8%; Score 35; DB 15; Length 162;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8
Db 18 LSYDLFV 24

RESULT 12
US-09-809-391-462
; Sequence 462, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 462
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (163)
; OTHER INFORMATION: Xaa equals stop translation
US-09-809-391-462

Query Match          77.8%; Score 35; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8
Db 18 LSYDLFV 24

RESULT 13
US-09-882-171-462
; Sequence 462, Application US/09882171
; Publication No. US20030175858A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
```





; PRIOR APPLICATION NUMBER: 60/047,501  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 60/043,670  
; PRIOR FILING DATE: 1997-04-11  
; PRIOR APPLICATION NUMBER: 60/056,632  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,664  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,876  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,881  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,909  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,875  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,862  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,887  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/056,908  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/048,964  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/057,650  
; PRIOR FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: 60/056,884  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 60/057,669  
; PRIOR FILING DATE: 1997-09-05

Query Match 77.8%; Score 35; DB 10; Length 163;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
| | | | |  
Db 18 LSYDLFV 24

RESULT 14  
US-10-164-861-462  
; Sequence 462, Application US/10164861  
; Publication No. US20030225248A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/10/164,861  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US/09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493  
; PRIOR FILING DATE: 1998-03-06  
; NUMBER OF SEQ ID NOS: 757  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 462  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (163)  
; OTHER INFORMATION: Xaa equals stop translation  
US-10-164-861-462

Query Match 77.8%; Score 35; DB 14; Length 163;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
| | | | |  
Db 18 LSYDLFV 24

RESULT 15  
US-09-867-550-916  
; Sequence 916, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Lesch, Martin D.  
; APPLICANT: Mehrahan, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 916  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-916

Query Match 75.6%; Score 34; DB 9; Length 55;  
Best Local Similarity 77.8%; Pred. No. 48;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
| | | | |  
Db 16 FLSYDLFV 24

Search completed: December 30, 2004, 21:00:34  
Job time : 9.81864 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:25:59 ; Search time 10.4962 Seconds  
(without alignments)  
307.593 Million cell updates/sec

Title: US-10-017-327-9  
Perfect score: 45  
Sequence: 1 FLSYDLFVV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	5	ABB08364
2	45	100.0	9	7	ABR82216
3	45	100.0	352	5	ABB08369
4	36	80.0	51	4	AAE74753
5	36	80.0	51	5	ABG65310
6	36	80.0	51	8	ADL78577
7	35	77.8	65	2	AAV13169
8	35	77.8	85	3	AAAG01223
9	35	77.8	130	2	AAV17226
10	35	77.8	130	4	AAU38996
11	35	77.8	130	5	ABE55705
12	35	77.8	160	2	AAV36168
13	35	77.8	162	2	AAV36215
14	35	77.8	162	3	AAV57894
15	35	77.8	162	4	AAH88605
16	35	77.8	162	5	ABH89619
17	35	77.8	162	5	ABP65087
18	35	77.8	162	7	ADD19028
19	35	77.8	162	7	ADJ45973
20	35	77.8	162	7	ADJ46067
21	35	77.8	162	8	ADP19476
22	35	77.8	163	2	AAW74871
23	35	77.8	163	5	ABG95322
24	35	77.8	163	6	ABO34516
25	35	77.8	163	7	ADI23177

26	35	77.8	163	8	ADH74179	Adh74179 Human sec
27	35	77.8	177	5	ABB97496	Abb97496 Novel hum
28	35	77.8	203	5	AAE25673	Aae25673 Bacillus
29	35	77.8	203	7	ABR62959	Abt62959 BstVI res
30	34	75.6	51	4	AAO08931	Aao08931 Human pol
31	34	75.6	55	5	ABP64088	Abp64088 Human ORF
32	34	75.6	98	4	AAG73909	Aag73909 Human col
33	34	75.6	160	7	ADC32867	Adc32867 Human nov
34	34	75.6	247	7	ADM04330	Adm04330 Human pro
35	34	75.6	363	6	ADA34406	Ada34406 Acinetoba
36	34	75.6	412	2	AAW98765	Aaw98765 H. pylori
37	34	75.6	516	5	ABB93691	Abb93691 Herbicida
38	34	75.6	576	7	ADM05144	Adm05144 Human pro
39	34	75.6	674	6	ADA55118	Ada55118 Human pro
40	34	75.6	1031	8	ADM72293	Adm72293 Equine TL
41	33	73.3	90	3	AAG38087	Aag38087 Arabidops
42	33	73.3	103	7	ADF04479	Adf04479 Bacterial
43	33	73.3	119	3	AAAG14935	Aag14935 Arabidops
44	33	73.3	181	3	AAG38086	Aag38086 Arabidops
45	33	73.3	211	5	ADH32233	Adh32233 Novel yea

ALIGNMENTS

RESULT 1  
ABB08364  
ID ABB08364 standard; protein; 9 AA.

AC ABB08364;

DT 07-MAY-2002 (first entry)

DE Synthetic epitope 4 of human cancer antigen eIF3.

XX Human; melanoma antigen eukaryotic initiation factor 3; eIF3;

KW ovarian cancer; MHC; cytosolic; immunomodulator; immune effector cell;

KW anti-cancer; vaccine.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1 /note= "HLA-2 binding residue"

FT Domain 2 /note= "HLA-2 binding residue"

FT Domain 3 /note= "T-cell receptor (TCR) binding domain"

FT Domain 9 /note= "HLA-2 binding residue"

WO200192307-A2.

06-DEC-2001.

30-MAY-2001; 2001WO-US017456.

31-MAY-2000; 2000US-0209391P.

17-AUG-2000; 2000US-0226258P.

20-DEC-2000; 2000US-0257008P.

(GENZ ) GENZYME CORP.

Nicolette CA;

WPI; 2002-139606/18.

N-PSDB; ABA97215.

New therapeutic compounds useful against human ovarian cancer, for modulating immune response in a subject, and for generating antibodies that specifically recognize and bind to these molecules.

Claim 29; Page 59; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen eIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents synthetic epitope 4 of human cancer antigen eIF3  
 XX Sequence 9 AA;

Query Match 100.0%; Score 45; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFVV 9  
 Db 1 FLSYDLFVV 9  
 |||||

RESULT 2  
 ABR82216  
 ID ABR82216 standard; peptide; 9 AA.  
 XX  
 AC ABR82216;  
 DT 13-OCT-2003 (first entry)  
 DE Human antigen eIF3 derived compound 4.  
 DE Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
 KW cytostatic; gene therapy; human; antigen.  
 XX Synthetic.  
 OS Homo sapiens.  
 OS  
 PN WO2003050543-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 05-DEC-2001; 2001WO-US047997.  
 XX  
 PR 05-DEC-2001; 2001WO-US047997.  
 XX  
 PA (GENZ ) GENZYME CORP.  
 XX  
 FI Nicolette CA;  
 XX  
 DR WPI; 2003-532936/50.  
 DR N-PSDB; ACC85033.  
 XX  
 XX Aiding in the diagnosis of a neoplastic condition, useful for treating  
 FT cancer and related malignancies comprises determining the amount of  
 FT expression of an eIF3 protein in a test sample isolated from the cell or  
 FT tissue.  
 FT  
 PS Claim 12; Page 30; 77pp; English.  
 XX  
 CC The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic

CC condition or susceptibility to a neoplastic condition based on the amount  
 CC of expression of the eIF3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. Sequences ABR82213-16  
 CC represent compounds derived from the human antigen eIF3  
 XX Sequence 9 AA;

Query Match 100.0%; Score 45; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFVV 9  
 Db 1 FLSYDLFVV 9  
 |||||

RESULT 3  
 ABB08369  
 ID ABB08369 standard; protein; 352 AA.  
 XX  
 AC ABB08369;  
 DT 07-MAY-2002 (first entry)  
 DE Human cancer antigen eIF3 variant 4 amino acid sequence.  
 DE Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 242 /note= "wild-type Asn is replaced by Phe"  
 FT Domain 242 /note= "HLA-2 binding residue"  
 FT Domain 243 /note= "HLA-2 binding residue"  
 FT Domain 244.249 /note= "T-cell receptor (TCR) binding domain"  
 FT Misc-difference 244 /note= "wild-type Gln is replaced by Ser"  
 FT Misc-difference 245 /note= "wild-type Leu is replaced by Tyr"  
 FT Misc-difference 246 /note= "wild-type Leu is replaced by Asp"  
 FT Misc-difference 247 /note= "wild-type Met is replaced by Leu"  
 FT Misc-difference 248 /note= "wild-type Asp is replaced by Phe"  
 FT Misc-difference 249 /note= "wild-type Arg is replaced by Val"  
 FT Domain 250 /note= "HLA-2 binding residue"  
 FT  
 XX WO200192307-A2.  
 PN  
 PD 06-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US017456.  
 XX  
 PR 31-MAY-2000; 2000US-0209391P.  
 PR 17-AUG-2000; 2000US-0226258P.  
 PR 20-DEC-2000; 2000US-0257008P.  
 XX  
 XX (GENZ ) GENZYME CORP.  
 PA  
 XX Nicolette CA;  
 PI  
 DR WPI; 2002-139606/18.

XX New therapeutic compounds useful against human ovarian cancer, for  
PT modulating immune response in a subject, and for generating antibodies  
PT that specifically recognize and bind to these molecules.  
XX  
XX Claim 9; Page; 69pp; English.  
XX  
XX The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory  
CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.  
CC Compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterised by expression of antigen E1F3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detecting genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents the human cancer antigen e1f3 variant 4 amino acid sequence.  
CC Note: This sequence is not present in the specification, but may be  
CC created from the sequence of the wild-type human cancer antigen e1f3  
CC sequence given in ABB08360  
XX  
XX Sequence 352 AA;  
SQ

Query Match 100.0%; Score 45; DB 5; Length 352;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
Db 242 FLSYDLFVV 250  
|||||

RESULT 4  
AAB74753  
ID AAB74753 standard; protein; 51 AA.  
XX  
XX AAB74753;  
AC  
XX  
XX 12-JUN-2001 (first entry)  
XX  
XX Human secreted protein sequence encoded by gene 21 SEQ ID NO:62.  
XX  
XX Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;  
XX dermatological; immunosuppressive; antiinflammatory; anti-HIV;  
XX immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;  
XX ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;  
XX antialzheimer; antiparkinsonian; antimicrobial; vulnery; gene therapy;  
XX immune disorder; hyperproliferative disorder; cardiovascular disease;  
XX cancer; angiogenic disorder; neurological disorder; infectious disease;  
XX wound healing; regeneration; chemotaxis.  
XX  
XX Homo sapiens.  
XX  
XX WO200112775-A2.  
XX  
XX 22-FEB-2001.  
XX  
XX 16-AUG-2000; 2000WO-US022325.  
XX  
XX 17-AUG-1999; 99US-0149182P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;  
PI Birse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;  
XX

DR WPI; 2001-147550/15.  
DR N-PSDB; AAF81807.  
XX  
XX Nucleic acids encoding 25 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
PT and diabetic retinopathy.  
XX  
XX Claim 11; Page 471; 485pp; English.  
XX  
XX AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733  
CC to AAB74772. Human secreted proteins can have activities based on the  
CC tissues and cells they are expressed in. Example of activities include:  
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;  
CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;  
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;  
CC anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and  
CC vulnery. Human secreted proteins can be used in gene therapy and  
CC vaccine. Human secreted protein nucleotide sequences (NAMI) and proteins  
CC (PEPI) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. For example, NAMI  
CC and PEPI may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patients genome that  
CC affect the activity of proteins by expressing inactive proteins or to  
CC supplement the patients own production of polypeptides. Disorders that  
CC may be prevented, diagnosed and/or treated include immune disorders,  
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,  
CC angiogenic disorders, neurological disorders, infectious diseases and/or  
CC for promoting wound healing, regeneration and /or chemotaxis. AAF81778 to  
CC AAF81786 and AAB74732 represent sequences used in the exemplification of  
CC the present invention  
XX  
XX Sequence 51 AA;  
SQ

Query Match 80.0%; Score 36; DB 4; Length 51;  
Best Local Similarity 66.7%; Pred. No. 7.1;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
Db 7 FVSYDYFIV 15  
|:|:|:|:|

RESULT 5  
ABG65310  
ID ABG65310 standard; protein; 51 AA.  
XX  
XX ABG65310;  
AC  
XX  
XX 27-AUG-2002 (first entry)  
XX  
XX Human albumin fusion protein #1985.  
XX  
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;  
XX human serum albumin; HSA; cancer; reproductive disorder;  
XX digestive disorder; immune disorder; endocrine disorder;  
XX haematopoietic disorder; neural disorder; connective disorder;  
XX cytostatic; antiinfertility; antiinflammatory; antiulcer;  
XX immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;  
XX neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
XX osteopathic; antiarthritic.  
XX  
XX Homo sapiens.  
XX Synthetic.  
XX  
XX WO200177137-A1.  
XX  
XX 18-OCT-2001.  
XX  
XX 12-APR-2001; 2001WO-US011988.  
XX  
XX 12-APR-2000; 2000US-0229358P.  
XX 25-APR-2000; 2000US-0199384P.  
XX 21-DEC-2000; 2000US-0256931P.  
XX



XX DE Human secreted protein encoded by 5' EST SEQ ID NO: 183.  
 XX DE  
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 KW forensic; gene therapy; chromosome mapping; signal peptide;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition.  
 XX KW  
 XX Homo sapiens.  
 OS XX  
 XX WO9906552-A2.  
 FN XX  
 XX 11-FEB-1999.  
 PD XX  
 XX 31-JUL-1998; 98WO-IB001236.  
 PF XX  
 XX 01-AUG-1997; 97US-00905223.  
 PR XX  
 XX (GEST ) GENSET.  
 PA XX  
 XX Dumas Milne Edwards J, Duclert A, Lacroix B;  
 PI XX  
 XX WPI; 1999-153782/13.  
 DR XX  
 DR N-PSDB; AAX51969.  
 XX  
 XX New isolated brain-derived nucleic acids - used to develop products which  
 PT may have cytokine, immune, regulatory, haematopoiesis regulating, anti-  
 PT inflammatory or tumour inhibition activity.  
 PT  
 PS Claim 34; Page 550; 577pp; English.  
 XX  
 XX AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAY12987 to  
 CC AAY13219, respectively. The proteins given represent the signal peptide  
 CC and an N-terminal fragment of a secreted protein. The nucleic acid  
 CC sequences can be used for producing secreted human gene products. They  
 CC can also be used to develop products for diagnosis and therapy. The  
 CC proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductive hormone  
 CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter,  
 CC sequences. The nucleic acids encoding the signal peptide can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 XX polypeptide into a membrane, or importing a polypeptide into a cell  
 XX  
 XX Sequence 65 AA;  
 SQ  
 Query Match 77.8%; Score 35; DB 2; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 LSYDLFV 8  
 Db |||||  
 18 LSYDLFV 24  
 RESULT 8  
 AAG01223  
 ID AAG01223 standard; protein; 85 AA.  
 XX AC  
 XX AAG01223;  
 XX  
 XX 06-OCT-2000 (first entry)  
 DT  
 DE Human secreted protein, SEQ ID NO: 5304.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping.  
 XX  
 XX Homo sapiens.  
 XX  
 XX EP1033401-A2.  
 FN  
 XX 06-SEP-2000.  
 PD  
 XX 21-FEB-2000; 2000EP-00200610.  
 PF  
 XX 26-FEB-1999; 99US-0122487P.  
 PR  
 XX (GEST ) GENSET.  
 XX  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI  
 XX WPI; 2000-500381/45.  
 DR  
 DR N-PSDB; AAC01229.  
 XX  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 PT  
 XX Claim 13; SEQ ID NO 5304; 71pp + Sequence Listing; English.  
 PS  
 XX The present sequence is a polypeptide encoded by one of a large number of  
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
 CC sequences derived from the 5' ends of mRNAs and even in those cases where  
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
 CC are also used in diagnostic, forensic, gene therapy and chromosome  
 CC mapping procedures. They are used to obtain upstream regulatory sequences  
 CC and to design expression and secretion vectors  
 XX  
 XX Sequence 85 AA;  
 SQ  
 Query Match 77.8%; Score 35; DB 3; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 LSYDLFV 8  
 Db |||||  
 18 LSYDLFV 24  
 RESULT 9  
 AAY17226  
 ID AAY17226 standard; protein; 130 AA.  
 XX AC  
 XX AAY17226;  
 XX  
 XX 09-AUG-1999 (first entry)  
 DT  
 XX Human secreted protein (clone pe204-1).  
 DE  
 XX Secreted protein; kidney; lung; brain; blood; testis; bone marrow;  
 XX nutritional activity; cytokine; cell proliferation; immune stimulation;  
 KW hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;  
 KW anti-inflammatory; tumour invasion.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO9926961-A1.  
 FN  
 XX 03-JUN-1999.  
 PD  
 XX 24-NOV-1998; 98WO-US025149.  
 XX  
 XX

```

PR 26-NOV-1997; 97US-0066804P.
XX 23-NOV-1998; 98US-00137886.
XX (GEMY ) GENETICS INST INC.
XX
PI Jacobs K, McCooy JM, Lavallie E, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Wong GG, Clark HF;
PI Fechtel K;
XX
DR WPI; 1999-357809/30.
DR N-PSDB; AAX60808.
XX
XX New polynucleotides encoding secreted proteins.
XX
XX Claim 34; Page 123; 133pp; English.
XX
XX The invention relates to secreted proteins (AA17219-228) encoded by
CC polynucleotides obtained from human fetal kidney, adult lung, adult
CC kidney, adult brain, adult blood, adult testes, and fetal brain and
CC murine adult bone marrow cDNA libraries. The secreted protein nucleic acid
CC sequences (X6801-811) correspond to clones bd306-7, fj283-6, fk317-3,
CC k213-2x, na316-1, nf93-20, np164-1, pe204-1, yal-1 and yb-1, (all clones
CC are deposited as ATCC 9599); The PNs and proteins are predicted to have
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and animals,
CC although no supporting data is given. Suggested activities include
CC nutritional activity, cytokine and cell proliferation/differentiation
CC activity, immune stimulating (e.g. as vaccines) or suppressing activity,
CC hematopoiesis regulating activity, tissue growth activity, activin/
CC inhibin activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The PNs are also stated to be useful for gene
CC therapy
XX
XX Sequence 130 AA;
XX
XX Query Match 77.8%; Score 35; DB 2; Length 130;
XX Best Local Similarity 100.0%; Pred. No. 31;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 LSYDLFV 8
XX | | | | |
XX Db 18 LSYDLFV 24
XX
XX RESULT 10
XX AAU38996
XX ID AAU38996 standard; protein; 130 AA.
XX
XX AC AAU38996;
XX
XX XX 16-JAN-2002 (first entry)
XX
XX DE Human secreted protein pe241_1.
XX
XX Human; secreted protein; antiinflammatory; immunosuppressive; nootropic;
XX neuroprotective; antiarthritic; antimicrobial; vulnary; cyostatic;
XX antidiabetic; virucide; antiinfertility; anticonvulsant; vasotropic;
XX antiparkinsonian; immunostimulant; dermatological; antirheumatic;
XX antitumor; antiulcer; osteoporosis; tranquiliser; cerebroprotective;
XX cytokine; cell proliferation; cell differentiation; immune deficiency;
XX severe combined immunodeficiency; SCID; tumour; autoimmune disorder;
XX multiple sclerosis; rheumatoid arthritis; graft-versus-host disease;
XX myeloid deficiency; wound healing; ulcer; periodontal disease;
XX osteoporosis; osteoarthritis; Alzheimer's disease; Parkinson's disease;
XX Huntington's disease; infection; cardiac disease; stroke; sepsis;
XX inflammatory bowel disease; contraceptive; immunogen; food supplement;
XX vaccine.
XX
XX Homo sapiens.
XX
XX OS
XX
XX PN W0200175068-A2.

```

---

```

XX PD 11-OCT-2001.
XX
XX 22-MAR-2001; 2001WO-US009369.
XX
XX 30-MAR-2000; 2000US-00539330.
XX 04-DEC-2000; 2000US-00729674.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Jacobs K, McCooy JM, Lavallie E, Collins-Racie LA, Evans C;
XX Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG, Clark H;
XX Fechtel K, Merberg D;
XX
XX WPI; 2001-639363/73.
XX N-PSDB; AAS59214.
XX
XX Secreted human proteins, useful as vaccine for treating various diseases
XX such as autoimmune disorders (e.g. multiple sclerosis), and nervous
XX system disorders (e.g. stroke).
XX
XX Disclosure; Page 468-469; 619pp; English.
XX
XX The invention relates to novel human secreted proteins, the nucleic acids
XX encoding them. The protein may exhibit cytokine, cell proliferation or
XX cell differentiation activity or may induce production of other cytokines
XX in certain cell populations and may exhibit immune stimulating or immune
XX suppressing activity, which is useful for the treatment of various immune
XX deficiencies and disorders e.g. severe combined immunodeficiency (SCID),
XX autoimmune disorders e.g. multiple sclerosis, systemic lupus
XX erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation.
XX The proteins are also useful in the treatment of diseases and disorders
XX including tissue, skin and organ transplantation and in graft-versus-host
XX diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid
XX cell deficiencies, wound healing and tissue repair, in the treatment of
XX burns, incisions and ulcers; as well as in treatment of periodontal
XX disease, osteoporosis or osteoarthritis, mediated by inflammatory
XX processes, diseases of the peripheral nervous system, Alzheimer's,
XX Parkinson's disease, Huntington's disease, amyotrophic lateral
XX sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and
XX central nervous system vessel e.g. stroke, sepsis, inflammatory bowel
XX disease, ulcers, bone regeneration. The protein, having activin- or
XX inhibin-related activities is useful as a contraceptive based on the
XX ability of inhibits to decrease fertility in female mammals and decrease
XX spermatogenesis in male mammals. The proteins and nucleic acids are also
XX useful as food supplements. The present sequence represents a secreted
XX protein of the invention
XX
XX Sequence 130 AA;
XX
XX Query Match 77.8%; Score 35; DB 4; Length 130;
XX Best Local Similarity 100.0%; Pred. No. 31;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 LSYDLFV 8
XX | | | | |
XX Db 18 LSYDLFV 24
XX
XX RESULT 11
XX ABB55705
XX ID ABB55705 standard; protein; 130 AA.
XX
XX AC ABB55705;
XX
XX XX 14-FEB-2002 (first entry)
XX
XX DE Human polypeptide SEQ ID NO 16.
XX
XX Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
XX immune disorder; bacterial infection; fungal infection; cancer; tumour;
XX autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
XX osteoporosis; osteoarthritis; nervous system disorder; neuropathy;

```



KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;  
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulneryary;  
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;  
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;  
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.  
 XX Homo sapiens.  
 OS  
 XX  
 XX US2001039335-A1.  
 XX  
 XX PD 08-NOV-2001.  
 XX  
 XX PF 04-DEC-2000; 2000US-00729674.  
 XX  
 XX PR 26-NOV-1997; 97US-0126425P.  
 XX PR 04-DEC-1997; 97US-0067454P.  
 XX PR 20-DEC-1997; 97US-0068379P.  
 XX PR 02-JAN-1998; 98US-0070348P.  
 XX PR 07-JAN-1998; 98US-0070643P.  
 XX PR 08-JAN-1998; 98US-0070755P.  
 XX PR 13-JAN-1998; 98US-0071304P.  
 XX PR 22-JAN-1998; 98US-0072134P.  
 XX PR 30-JAN-1998; 98US-0073095P.  
 XX PR 18-FEB-1998; 98US-0075038P.  
 XX PR 23-NOV-1998; 98US-00197886.  
 XX PR 30-MAR-2000; 2000US-00539330.  
 XX  
 XX PA (JACO/) JACOBS K.  
 XX PA (MCCO/) MCCOY J M.  
 XX PA (LAVA/) LAVALLIE E R.  
 XX PA (COLL/) COLLINS-RACIE L A.  
 XX PA (EVAN/) EVANS C.  
 XX PA (MERB/) MERBERG D.  
 XX PA (TREA/) TREACY M.  
 XX PA (AGOS/) AGOSTINO M J.  
 XX PA (STEI/) STEININGER R J.  
 XX PA (SPAU/) SPAULDING V.  
 XX PA (WONG/) WONG G G.  
 XX PA (CLAR/) CLARK H.  
 XX PA (FECH/) FECHTEL K.  
 XX  
 XX PI Jacobs K, McCooy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 XX PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;  
 XX PI Wong GG, Clark H, Fecht K;  
 XX  
 XX WPI; 2002-040725/05.  
 XX N-PSDB; ABA90883.  
 XX  
 XX PT New secreted proteins and encoding polynucleotides, useful in gene  
 XX PT therapies, particularly for preventing or treating autoimmune disorders,  
 XX PT cancer, graft-versus-host disease, wound, osteoporosis, stroke or  
 XX PT inflammations.  
 XX  
 XX PS Disclosure; Page 189-190; 349pp; English.  
 XX  
 XX CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and  
 XX CC ABA90980) and encoded proteins (AB55698-AB55800), especially  
 XX CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and  
 XX CC proteins SEQ ID NO 2 (AB55698) and SEQ ID NO 20 (AB55707) contained in  
 XX CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1  
 XX CC are deposited with the American Type Culture Collection (ATCC) with  
 XX CC accession number 98599. The polynucleotides and encoded polypeptides have  
 XX CC cytostatic, anti-inflammatory, immunomodulator, vulneryary,  
 XX CC neuroprotective, activin, inhibit, chemotactic, haemostatic, thrombolytic  
 XX CC and anti-inflammatory activity and acting as cytokine modulators,  
 XX CC haematopoiesis regulators, tissue growth modulators and/or cadherin  
 XX CC suppressors. The polypeptides and polynucleotides are useful in gene  
 XX CC therapies, particularly for preventing, treating or ameliorating any of  
 XX CC the following diseases: immune deficiency and disorders; e.g. bacterial  
 XX CC or fungal infections, autoimmune disorders, cancer, systemic lupus  
 XX CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell  
 XX CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or  
 XX CC osteoarthritis; central and peripheral nervous system diseases and

CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's  
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;  
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis  
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,  
 CC endotoxin lethality, arthritis, inflammation, bowel disease or Crohn's  
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus  
 XX  
 XX SQ Sequence 130 AA;  
 Query Match 77.8%; Score 35; DB 5; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LSVDLFV 8  
 DB 18 LSVDLFV 24  
 |||||  
 |||||  
 RESULT 12  
 AAY36168  
 ID AAY36168 standard; protein; 160 AA.  
 XX  
 AC AAY36168;  
 XX  
 DT 23-SEP-1999 (first entry)  
 XX  
 DE Human secreted protein #40.  
 XX  
 KW Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9925825-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PF 13-NOV-1998; 98WO-IB001862.  
 XX  
 PR 13-NOV-1997; 97US-0066677P.  
 PR 17-DEC-1997; 97US-0069957P.  
 PR 09-FEB-1998; 98US-0074121P.  
 PR 13-APR-1998; 98US-0081563P.  
 PR 10-AUG-1998; 98US-0096116P.  
 PR 04-SEP-1998; 98US-0099273P.  
 XX  
 XX (GEST ) GENSET.  
 XX  
 XX Bougueleret L, Duclert A, Dumas Milne Edwards J;  
 XX WPI; 1999-347472/29.  
 XX N-PSDB; AAX97852.  
 XX  
 XX PT Extended cDNAs encoding secreted proteins.  
 XX  
 XX PS Example 28; Page 247-248; 307pp; English.  
 XX  
 XX CC AAY36129-Y36222 represent novel human secreted proteins encoded by the  
 XX CC extended cDNA sequences represented in AAX97813-X97906. The proteins of  
 XX CC the invention have cytostatic, thrombotic and osteopathic activity. The  
 XX CC extended cDNAs can be used to express secreted proteins or parts of them  
 XX CC or to obtain antibodies capable of binding to the secreted proteins. They  
 XX CC may also be used in diagnostic, forensic, gene therapy and chromosome  
 XX CC mapping procedures. Uses also include design of expression vectors and  
 XX CC secretion vectors  
 XX  
 XX SQ Sequence 160 AA;  
 Query Match 77.8%; Score 35; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LSVDLFV 8

```

Db      18 LSYDLFV 24
|||||
RESULT 13
AAY36215
ID AAY36215 standard; protein; 162 AA.
XX AC AAY36215;
XX AC AAY36215;
XX AC AAY36215;
DT 23-SEP-1999 (first entry)
XX Human secreted protein #87.
DE Human secreted protein #87.
XX Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;
KW diagnostic; gene therapy; chromosome mapping; secretion vector.
XX Homo sapiens.
XX WO925825-A2.
XX 27-MAY-1999.
XX 13-NOV-1998; 98WO-IB001862.
XX 13-NOV-1997; 97US-0066677P.
XX 17-DEC-1997; 97US-0069957P.
XX 09-FEB-1998; 98US-0074121P.
XX 13-APR-1998; 98US-0081563P.
XX 10-AUG-1998; 98US-0096118P.
XX 04-SEP-1998; 98US-0099273P.
XX (GEST ) GENSET.
XX Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX WPI; 1999-347472/29.
XX N-PSDB; AAX97899.
XX Extended cDNAs encoding secreted proteins.
XX Claim 7; Page 304-305; 307pp; English.
XX AAY36129-V36222 represent novel human secreted proteins encoded by the
CC extended cDNA sequences represented in AAX97813-X97906. The proteins of
CC the invention have cytostatic, thrombotic and osteopathic activity. The
CC extended cDNAs can be used to express secreted proteins or parts of them
CC or to obtain antibodies capable of binding to the secreted proteins. They
CC may also be used in diagnostic, forensic, gene therapy and chromosome
CC mapping procedures. Uses also include design of expression vectors and
CC secretion vectors
XX Sequence 162 AA;
Query Match 77.8%; Score 35; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LSYDLFV 8
Db 18 LSYDLFV 24
|||||
RESULT 14
AAY57894
ID AAY57894 standard; protein; 162 AA.
XX AC AAY57894;
XX AC AAY57894;
DT 23-MAR-2000 (first entry)
XX Human transmembrane protein HTPMPN-18.
DE Human transmembrane protein HTPMPN-18.
XX
Human; transmembrane protein; HTPMPN; diagnosis; immunospecific;
antiproliferative; neuroprotective; immune disorder;
reproductive disorder; smooth muscle disorder; neurological disorder;
gastrointestinal disorder; developmental disorder;
cell proliferative disorder.
XX Homo sapiens.
XX WO9961471-A2.
XX 02-DEC-1999.
XX 28-MAY-1999; 99WO-US011904.
XX 29-MAY-1998; 98US-0087260P.
XX 02-JUL-1998; 98US-0091674P.
XX 02-OCT-1998; 98US-0102954P.
XX 24-NOV-1998; 98US-0109869P.
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;
PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR, Au-Young J;
XX WPI; 2000-072605/06.
XX N-PSDB; AAZ56715.
XX Proteins, polynucleotides, vectors, host cells and antibodies used to
PT diagnose, treat or prevent immune, reproductive, smooth muscle,
PT neurological, gastrointestinal, developmental and cell proliferative
PT disorders.
XX Claim 1; Page 125; 229pp; English.
XX AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human
CC transmembrane proteins designated HTPMPN-1 to HTPMPN-79, respectively. The
CC transmembrane protein have immunospecific, antiproliferative and
CC neuroprotective activities. The human transmembrane proteins,
CC polynucleotides encoding them and other compositions and methods from the
CC present invention, can be used for the diagnosis, treatment or prevention
CC of immune, reproductive, smooth muscle, neurological, gastrointestinal,
CC developmental and cell proliferative disorders. The HTPMPN's can be used
CC to treat or prevent disorders associated with a decreased expression or
CC activity of HTPMPN
XX Sequence 162 AA;
Query Match 77.8%; Score 35; DB 3; Length 162;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LSYDLFV 8
Db 18 LSYDLFV 24
|||||
RESULT 15
AAB88605
ID AAB88605 standard; protein; 162 AA.
XX AC AAB88605;
XX AC AAB88605;
XX 04-JUN-2001 (first entry)
XX Human hydrophobic domain containing protein clone HP10773 #129.
XX Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW antianemic; vulnery; antilucer; osteopathic; anti-inflammatory;
KW cytostatic; gene therapy; autoimmune disorder; multiple sclerosis;
KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW behavioural characteristic; immune response.
XX

```

```

OS Homo sapiens.
XX
FN WO200112660-A2.
XX
XX
PD 22-FEB-2001.
XX
XX
PF 10-AUG-2000; 2000WO-JP0053356.
XX
XX
PR 17-AUG-1999; 99JP-00230344.
PR 07-SEP-1999; 99JP-00252551.
PR 01-OCT-1999; 99JP-00281112.
PR 22-OCT-1999; 99JP-00301624.
PR 04-NOV-1999; 99JP-00313877.
XX
XX (SAGA ) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
XX
XX WPI; 2001-160059/16.
DR N-PSDB; AAF94505.
XX
XX Human proteins with hydrophobic domains and the DNAs which encode them
PT are useful for treating autoimmune disorders, burns and tumors and for
PT screening novel pharmaceuticals.
XX
XX Claim 1; Page 478-479; 518pp; English.
XX
XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
CC AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,
CC anti-HIV, neuroprotective, antianaemic, vulnerary, antitumor,
CC osteoprotic, anti-inflammatory and cytostatic activities, and can be used
CC in gene therapy. (I) can be used as pharmaceuticals and as antigens to
CC prepare antibodies. DNA and cDNA (II) encoding (I) can be used as probes
CC for genetic diagnosis and gene sources for gene therapy or for producing
CC (I) in large quantities. Cells containing (II) are used for the detection
CC of ligands or receptors corresponding to membrane or secretory proteins
CC and to screen small molecule novel pharmaceuticals. Antibodies directed
CC to (I) can be used for the detection, quantification and purification of
CC (II). Activities of (I) may include cytokine and cell
CC proliferation/differentiation function, immune stimulating or suppressing
CC activity, haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory
CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.
CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
CC inflammatory bowel disease and tumours. (I) and (II) can also be used for
CC wound healing, as nutritional sources or supplements e.g. as amino acid,
CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
CC processing and utilisation of dietary fat, protein, carbohydrate,
CC vitamins and minerals, to effect behavioural characteristics, to affect
CC appetite, and can act as antigens in vaccines to raise an immune response
CC to the protein or another material cross-reactive with the protein
XX
XX
SQ Sequence 162 AA;
Query Match 77.8%; Score 35; DB 4; Length 162;
Best Local Similarity 100.0%; Pred.No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LSYDLFV 8
Db 18 LSYDLFV 24
| | | | |
| | | | |

```

Search completed: December 30, 2004, 20:43:04  
 Job time : 13.4962 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 30, 2004, 20:30:54 ; Search time 1.97229 Seconds  
(without alignments)  
439.058 Million cell updates/sec

Title: US-10-017-327-9

Perfect score: 45

Sequence: 1 FLSYDLFV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	91.1	247	2	AG2921
2	37	82.2	259	2	C89964
3	36	80.0	191	2	F72277
4	34	75.6	95	2	A10315
5	34	75.6	412	2	C64712
6	34	75.6	412	2	D71803
7	34	75.6	452	2	C72295
8	33	73.3	70	2	S01213
9	33	73.3	203	2	I40517
10	33	73.3	274	2	T39166
11	33	73.3	305	2	G86852
12	33	73.3	333	2	T17712
13	33	73.3	370	2	T47504
14	33	73.3	370	2	S40469
15	33	73.3	559	2	S62503
16	33	73.3	572	2	T50404
17	33	73.3	583	2	T14275
18	33	73.3	849	2	T20422
19	32	71.1	137	2	A83852
20	32	71.1	288	2	I78556
21	32	71.1	309	2	G69796
22	32	71.1	346	2	B90243
23	32	71.1	397	2	B98293
24	32	71.1	397	2	AG2990
25	32	71.1	413	2	D86920
26	32	71.1	586	2	T19075
27	32	71.1	614	2	T05387
28	32	71.1	726	1	S73915
29	32	71.1	800	2	T38032
					protein phosphatas
					truncated transpos
					hypothetical prote
					conserved hypothet
					ubiquinol-cytochro
					hypothetical prote
					NADH2 dehydrogenas
					spaG protein - Bac
					inorganic phosphat
					oxidoreductase ysfj
					hypothetical prote
					mitogen-activated
					inorganic phosphat
					probable inorganic
					probable inorganic
					hypothetical prote
					hypothetical prote
					membrane glycoprot
					lactose permease h
					aminomethyltransfe
					mannonate dehydrat
					probable UDP-galac
					hypothetical prote
					virulence-associat
					hypothetical prote

## ALIGNMENTS

### RESULT 1

A82921

protein phosphatase U0215 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: A82921

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: A82870

A:Accession: A82921

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <GLA>

A:Cross-references: GB:AB002120; GB:AF222894; NID:G6899167; PIDN:AAF30623.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: ptcl; U0215

A:Genetic code: SGC3

Query Match 91.1%; Score 41; DB 2; Length 247;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8

Db 176 FLSYDLFV 183

### RESULT 2

C89964

truncated transposase [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C:Accession: C89964

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oqu

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; PMID:21311952; PMID:11418146

A:Accession: C89964

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-259 <KUR>

A:Cross-references: UNIPROT:Q99T82; GB:BA000018; PID:gi13701580; PIDN:BA042873.1; GSPDB:G

A:Experimental source: strain N315

C:Genetics:

A:Gene: truncated-SA

Query Match 82.2%; Score 37; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7  
 12 FLSYDLF 18

Db

RESULT 3  
 F72277  
 hypothetical protein - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: F72277  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
 C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72277

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-191 <ARN>

A:Cross-references: UNIPROT:Q9XOX9; GB:AE001780; GB:AE000512; NID:g4981796; PIDN:AAD3632

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1252

Query Match 80.0%; Score 36; DB 2; Length 191;  
 Best Local Similarity 66.7%; Pred. No. 8.5;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 44 FLSYDLFVI 52

Db

RESULT 4  
 A10315  
 conserved hypothetical protein YPO2589 [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: A10315  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Hoiden, M.T.G.; Prentice, M.B.  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
 Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB00001; MUID:21470413; PMID:11586360

A:Accession: A10315

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-95 <KUR>

A:Cross-references: UNIPROT:Q8ZDH9; GB:AL590842; PIDN:CAC91389.1; PID:g15980575; GSPDB:G

C:Genetics:

A:Gene: YPO2589

Query Match 75.6%; Score 34; DB 2; Length 95;  
 Best Local Similarity 85.7%; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 36 LSYDLFI 42

Db

RESULT 5  
 C64712  
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Helicobacter pylori (strain  
 C:Species: Helicobacter pylori  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
 C:Accession: C64712  
 R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
 Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: C64712

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-412 <TOM>

A:Cross-references: UNIPROT:O26064; GB:AE000652; GB:AE000511; NID:g2314720; PIDN:AAD0857;

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol-

C:Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase

F:21-370/Domain: cytochrome b homology <CBH>

F:21-223/Domain: cytochrome b6 homology <CB6>

F:250-370/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:94,195/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:108,210/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 6  
 D71803  
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Helicobacter pylori (strain  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: D71803  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
 Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: D71803

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-412 <ARN>

A:Cross-references: UNIPROT:Q9ZJ55; GB:AE001568; GB:AE001439; NID:g4156083; PIDN:AAD0704;

A:Experimental source: strain J99

C:Genetics:

A:Gene: petB

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol-

C:Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase

F:21-370/Domain: cytochrome b homology <CBH>

F:21-223/Domain: cytochrome b6 homology <CB6>

F:250-370/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:94,195/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:108,210/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 253 FLSYDLFV 261

Db

RESULT 7  
 C72295  
 hypothetical protein TM1105 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: C72295  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
 A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: C72295  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-452 <ARN>  
 A:Cross-references: UNIPROT:Q9X019; GB:AE001769; GB:AE000512; NID:g4981639; PIDN:AAD3618  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM1105  
 C:Superfamily: [NiFe]-hydrogenase-3-type complex, large membrane subunit

Query Match 75.6%; Score 34; DB 2; Length 452;  
 Best Local Similarity 85.7%; Pred. No. 49;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLF 7  
 Db 102 FVSYDLF 108  
 |:|||||

RESULT 8  
 S01213  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 3 - brine shrimp mitochondrion (frag  
 C:Species: mitochondrion Artemia sp. (brine shrimp)  
 C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 03-Jun-2002  
 C:Accession: S01213  
 R:Bauecas, B.; Garesse, R.; Calleja, M.; Valverde, J.R.; Marco, R.  
 Nucleic Acids Res. 16, 6515-6529, 1988  
 A:Title: Genome organization of Artemia mitochondrial DNA.  
 A:Reference number: S01207; MUID:88289417; PMID:3135541  
 A:Accession: S01213  
 A:Molecule type: DNA  
 A:Residues: 1-31; 32-70 <BAT>  
 A:Cross-references: EMBL:X07665  
 A:Note: the translation of residues 1-8 is inconsistent with the nucleotide sequence  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Genetic code: SGC4  
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3  
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 73.3%; Score 33; DB 2; Length 70;  
 Best Local Similarity 75.0%; Pred. No. 13;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8  
 Db 33 FLSYPLFI 40  
 |:|||||

RESULT 9  
 I40517  
 spaG protein - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
 C:Accession: I40517  
 R:Klein, C.; Entian, K.D.  
 Appl. Environ. Microbiol. 60, 2793-2801, 1994  
 A:Title: Genes involved in self-protection against the lantibiotic subtilin produced by  
 A:Reference number: I40511; MUID:94368094; PMID:8085823  
 A:Accession: I40517  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-203 <RES>  
 A:Cross-references: UNIPROT:Q45405; EMBL:U09819; NID:g2702240; PIDN:AAB91595.1; PID:g595  
 C:Genetics:  
 A:Gene: spaG

Query Match 73.3%; Score 33; DB 2; Length 203;  
 Best Local Similarity 62.5%; Pred. No. 36;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFVV 9  
 Db 78 LSYDIFIM 85  
 |:|||||

RESULT 10  
 T39166  
 inorganic phosphate transporter [similarity] - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
 C:Accession: T39166  
 R:Connor, R.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, February 1998  
 A:Reference number: 221832  
 A:Accession: T50381  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-274 <CO2>  
 A:Cross-references: EMBL:AL021815; PIDN:CAA16994.1; GSPDB:GN00067; SPDB:SPBC8E4.01c  
 A:Experimental source: strain 972h-; cosmid c8E4  
 C:Genetics:  
 A:Gene: SPDB:SPAC8E4.01c  
 A:Map position: 2  
 C:Superfamily: probable inorganic phosphate transport protein PHO84

Query Match 73.3%; Score 33; DB 2; Length 274;  
 Best Local Similarity 71.4%; Pred. No. 48;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYDLFVV 9  
 Db 59 SYDLFII 65  
 |:|||||

RESULT 11  
 G86852  
 oxidoreductase ysjB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
 C:Species: Lactococcus lactis subsp. lactis  
 C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: G86852  
 R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
 Genome Res. 11, 731-753, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
 A:Reference number: A86625; MUID:21235186; PMID:11337471  
 A:Accession: G86852  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-305 <STO>  
 A:Cross-references: UNIPROT:Q9CEL4; GB:AE005176; PID:g12724851; PIDN:AAK05921.1; GSPDB:G  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: ysjB

Query Match 73.3%; Score 33; DB 2; Length 305;  
 Best Local Similarity 62.5%; Pred. No. 53;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8  
 Db 155 FLMYDIFI 162  
 |:|||||

RESULT 12  
 T17712  
 hypothetical protein A222R - Chlorella virus PBCV-1  
 C:Species: Chlorella virus PBCV-1  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T17712  
 R:Graves, M.V.; Van Etten, J.L.  
 submitted to the EMBL Data Library, May 1999  
 A:Reference number: 218806

A;Accession: T17712  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A;Residues: 1-333 <GRA>  
A;Cross-references: UNIPROT:Q84542; EMBL:U42580; NID:G4028896; PIDN:AAC96590.1

A;Experimental source: specific host Chlorella strain NC64A  
C;Genetics:  
A;Note: A222R

C;Superfamily: Chlorella virus PBCV-1 hypothetical protein A222R

Query Match 73.3%; Score 33; DB 2; Length 333;  
Best Local Similarity 75.0%; Pred. No. 58;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8  
|:|||||

Db 254 FFSFDLFV 261  
|:|||||

## RESULT 13

T47504

mitogen-activated protein kinase 3 - Arabidopsis thaliana  
N;Alternate names: protein F9K21.220

C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C;Accession: T47504  
R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, February 2000  
A;Reference number: Z2467

A;Accession: T47504  
A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-370 <JOR>

A;Cross-references: UNIPROT:Q39023; EMBL:AL138657  
A;Experimental source: cultivar Columbia; BAC clone F9K21

C;Genetics:  
A;Map position: 3

A;Introns: 53/2; 96/3; 142/3; 253/3; 315/1  
A;Note: F9K21.220

C;Superfamily: kinase-related transforming protein; protein kinase homology  
Query Match 73.3%; Score 33; DB 2; Length 370;  
Best Local Similarity 71.4%; Pred. No. 64;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 7  
|:|||||

Db 21 FLSYDIF 27  
|:|||||

## RESULT 14

S40469

mitogen-activated protein kinase 3 (EC 2.7.1.1) - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 06-Oct-1994 #sequence\_revision 19-Jan-1996 #text\_change 09-Jul-2004  
C;Accession: S40469

R;Mizoguchi, T.; Hayashida, N.; Yamaguchi-Shinozaki, K.; Kamada, H.; Shinozaki, K.

FEBS Lett. 336, 440-444, 1993  
A;Title: ATPPKs: a gene family of plant MAP kinases in Arabidopsis thaliana.

A;Reference number: S40469; MUID:94109583; PMID:8282107  
A;Accession: S40469

A;Molecule type: mRNA  
A;Residues: 1-370 <MIZ>

A;Cross-references: UNIPROT:Q39023; EMBL:D21839; NID:G457397; PIDN:BAA04866.1; PID:G4573

C;Genetics:  
A;Gene: MPK3

C;Superfamily: kinase-related transforming protein; protein kinase homology  
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F;35-324/Domain: protein kinase homology <KIN>  
F;44-52/Region: protein kinase ATP-binding motif

Query Match 73.3%; Score 33; DB 2; Length 370;  
Best Local Similarity 71.4%; Pred. No. 64;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:29:59 ; Search time 10.5642 Seconds  
(without alignments)  
490.180 Million cell updates/sec

Title: US-10-017-327-9  
Perfect score: 45  
Sequence: 1 FLSYDLFV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	91.1	247	2 Q9PQS7	Q9Pqs7 ureaplasma
2	38	84.4	69	2 Q7BHL0	Q7bhl0 fusobacteri
3	38	84.4	69	2 AAS45303	Aas45303 fusobacte
4	37	82.2	259	2 Q99T82	Q99t82 staphylococ
5	36	80.0	191	2 Q9X0X9	Q9x0x9 thermotoga
6	35	77.8	101	2 Q9D9A8	Q9d9a8 mus musculu
7	35	77.8	162	2 Q9NPI0	Q9np10 homo sapien
8	35	77.8	162	2 Q9D6G5	Q9d6g5 m mus muscu
9	35	77.8	162	2 AAH58237	Aah58237 mus muscu
10	35	77.8	192	2 Q8T2F8	Q8t2f8 plasmodium
11	35	77.8	203	2 Q84AF2	Q84af2 bacillus st
12	34	75.6	35	2 Q88216	Q88216 pseudomonas
13	34	75.6	95	2 Q8ZDH9	Q8zdh9 yersinia pe
14	34	75.6	95	2 AAS61372	Aas61372 yersinia
15	34	75.6	184	2 Q6L299	Q6l299 picrophilus
16	34	75.6	382	2 Q9ZPF4	Q9zpf4 arabisopsis
17	34	75.6	389	1 UXUA_BACTN	Q8a7u2 bacteroides
18	34	75.6	394	1 UXUA_RHILO	Q987x5 rhizobium l
19	34	75.6	409	2 Q7VHG1	Q7vhg1 helicobacte
20	34	75.6	411	2 Q7M7Q1	Q7m7q1 wollinella s
21	34	75.6	412	2 Q26064	Q26064 helicobacte
22	34	75.6	412	2 Q9ZJ55	Q9zj55 helicobacte
23	34	75.6	452	2 Q9X0I9	Q9x0i9 thermotoga
24	34	75.6	516	2 Q9ZWT3	Q9zwt3 arabisopsis
25	34	75.6	590	2 Q8Y105	Q8y105 ralstonia s
26	34	75.6	618	2 Q7NIU9	Q7niu9 gloeobacter
27	34	75.6	674	1 KBT6_HUMAN	Q85v97 homo sapien
28	34	75.6	729	2 Q91L83	Q91l83 white spot
29	34	75.6	924	2 Q6CFR0	Q6cfro yarrowia li
30	34	75.6	1559	2 Q8IDC3	Q8idc3 plasmodium
31	33	73.3	70	1 NU3M_ARTSA	P19043 artemia sal

32	33	73.3	203	2 Q45405	Q45405 bacillus su
33	33	73.3	305	2 Q9CEL4	Q9cel4 lactococcus
34	33	73.3	332	2 Q72HT6	Q72ht6 thermus the
35	33	73.3	332	2 AAS81742	Aas81742 thermus t
36	33	73.3	355	2 Q8RD34	Q8rd34 thermoanaer
37	33	73.3	370	1 MPK3_ARATH	Q39023 arabisopsis
38	33	73.3	370	2 AAN15326	Aan15326 arabidops
39	33	73.3	432	2 Q84542	Q84542 paramecium
40	33	73.3	479	2 Q9DHK3	Q9dhk3 yaba-like d
41	33	73.3	559	1 YAEC_SCHPO	Q09852 schizosacch
42	33	73.3	572	1 YEN1_SCHPO	Q42885 schizosacch
43	33	73.3	573	1 YHD1_SCHPO	Q9p6j9 schizosacch
44	33	73.3	583	2 Q9Y7Q9	Q9y7q9 schizosacch
45	33	73.3	607	2 Q8A4J7	Q8a4j7 bacteroides

ALIGNMENTS

RESULT 1  
Q9PQS7 PRELIMINARY; PRT; 247 AA.  
AC Q9PQS7; 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Protein phosphatase.  
GN Name=ptcl; OrderedLocNames=UU215;  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 700970;  
EX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
RT urealyticum.";  
RL Nature 407:757-762(2000).  
DR EMBL; AE002120; AAF30623.1; -  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR001932; PP2C-like.  
DR Pfam; PF00481; PP2C; 1.  
DR SMART; SM00332; PP2Cc; 1.  
DR SMART; SM00331; PP2C\_SIG; 1.  
KW Complete proteome.  
SQ SEQUENCE 247 AA; 28322 MW; A81EFE16AF275C80 CRC64;  
Query Match 91.1%; Score 41; DB 2; Length 247;  
Best Local Similarity 100.0%; Pred.No. 6.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FLSYDLFV 8  
Db 176 FLSYDLFV 183  
RESULT 2  
Q7BHL0 PRELIMINARY; PRT; 69 AA.  
AC Q7BHL0; 05-JUL-2004 (TremBLrel. 27, Created)  
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)  
DE Oppf (Fragment).  
GN Name=oppf;  
OS Fusobacterium necrophorum.  
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
OC Fusobacterium.  
OX NCBI\_TaxID=859;  
RN [1]  
RP SEQUENCE FROM N.A.



```

Query Match      80.0%; Score 36; DB 2; Length 191;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLSYDLFVV 9
      |||:| |
Db      44 FLSFDFVVI 52

RESULT 6
Q9D9A8      PRELIMINARY;      PRT;      101 AA.
AC Q9D9A8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:1700113101 product:hypothetical protein, full insert
DE sequence.
GN Name=1700113101Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayateu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto S., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;

```

```

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyma T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AK007197; BAB24895.1; -.
DR MGI; MGI:1923914; 1700113101Rik.
DR GO; GO:0005085; P:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR000651; RasGef N.
DR PROSITE; PS50212; RASGEF_NTER; 1.
KW Hypothetical protein.
SQ SEQUENCE 101 AA; 11635 MW; E9277CDA50D73D30 CRC64;

Query Match      77.8%; Score 35; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LSYDLFV 8
      ||| |
Db      18 LSYDLFV 24

RESULT 7
Q9NP10      PRELIMINARY;      PRT;      162 AA.
AC Q9NP10;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE HSPC198 (Hypothetical protein HSPC196).
GN Name=HSPC196;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSU=Umbilical cord blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,
RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.,
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSU=Bladder;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

```

RA Krzywinski M.I., Skaleka U., Smalilus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bladder;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL; AF151032; AAF36118.1; -;  
 DR EMBL; BC005201; AAH05201.1; -;  
 DR EMBL; AF151030; AAF36116.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 162 AA; 19262 MW; 31D76DD5C953333DE CRC64;  
 Query Match 77.8%; Score 35; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LSYDLFV 8  
 DB 18 LSYDLFV 24  
 RESULT 8  
 Q9D6G5  
 ID Q9D6G5 PRELIMINARY; PRT; 162 AA.  
 AC Q9D6G5;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched  
 DE library, clone:2900055D14 product:hypothetical protein, full insert  
 DE sequence (Mus musculus adult pancreas islet cells cDNA, RIKEN full-  
 DE length enriched library, clone:C820001003 product:hypothetical  
 DE protein, full insert sequence) (RIKEN CDNA 2900055D14).  
 GN Name:2900055D14R1k;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]\_TaxID=10090;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44 (1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690 (2001).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata K., Hayashizaki Y., Sugahara Y., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Hatada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771 (2000).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus;  
 ADACHI J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa Y., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
 ADACHI J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.  
 RN [8]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altchul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [9]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RA Strausberg R.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBSJ databases.

```

DR EMBL; AK013694; BAB28958.1; -.
DR EMBL; AK050471; BAC34274.1; -.
DR EMBL; BC058237; AAH58237.1; -.
DR MGD; MGI:1920232; 2900055D14Rik.
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 19150 MW; 4E1D72AD014E29E0 CRC64;

Query Match          77.8%; Score 35; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8
   |||||
Db 18 LSYDLFV 24

RESULT 9
AAH58237 PRELIMINARY; PRT; 162 AA.
AC AAH58237;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE RIKEN CDNA 2900055D14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058237; AAH58237.1; -.
SQ SEQUENCE 162 AA; 19150 MW; 4E1D72AD014E29E0 CRC64;

Query Match          77.8%; Score 35; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8
   |||||
Db 18 LSYDLFV 24

RESULT 10
Q812F8 PRELIMINARY; PRT; 192 AA.
ID Q812F8
AC Q812F8;

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PF11750c.
GN Name=PF11750c;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagsels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL929359; CAD52036.1; -.
DR InterPro; IPR006496; DUF_Pfalc267.
DR TIGRFAMs; TIGR01609; PF_unchar_267; 1.
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 23573 MW; DCFMA5B8802CC7A3 CRC64;

Query Match          77.8%; Score 35; DB 2; Length 192;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9
   |||||
Db 37 FLSYDMLII 45

RESULT 11
Q84AF2 PRELIMINARY; PRT; 203 AA.
ID Q84AF2
AC Q84AF2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BstYI.
GN Name=BstYIR;
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y406;
RA Xu S.-Y., Samuelson J., Pelletier J., Sibley M., Wilson G.G.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY197779; AA048714.1; -.
SQ SEQUENCE 203 AA; 23193 MW; C6F5C22A35DB04EF CRC64;

Query Match          77.8%; Score 35; DB 2; Length 203;
Best Local Similarity 75.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8
   |||||
Db 136 FVAYDLFV 143

RESULT 12

```

Q88216 ID Q88216 PRELIMINARY; PRT; 35 AA.  
AC Q88216;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Lipoprotein, putative.  
GN OrderedLocusNames=PSPT02641;  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
RA Gwinn M.L., Dordon R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,  
RA Madupu R., Dougherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
RA Nelson W.C., Davidsten T.M., Zafar N., Zhou L., Liu J., Yuan Q.,  
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,  
RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,  
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,  
RA Bender C.L., White O., Fraser C.M., Collier A.;  
RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
Pseudomonas syringae pv. tomato DC3000."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).  
DR EMBL; AEO16865; AAC56144.1; -.  
DR TIGR; PSPT02641; -.  
KW Complete proteome; Lipoprotein.  
SQ SEQUENCE 35 AA; 3854 MW; E209450422A75C09 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 35;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFWV 9  
| | | | |  
Db 9 FLSYGLYVI 17

RESULT 13  
Q8ZDH9 ID Q8ZDH9 PRELIMINARY; PRT; 95 AA.  
AC Q8ZDH9; Q74VZ3; Q7CJW2;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein YP02589 (Hypothetical protein y1159).  
GN OrderedLocusNames=Yp1126, YP02589, y1159;  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIMS / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Retherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM."  
RL J. Bacteriol. 184:4601-4611(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91001 / Biovar Mediaevalis;  
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,  
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,  
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
RA Yang R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A414152; CAC91389.1; -.  
DR EMBL; AEO13719; AAM84736.1; -.  
DR EMBL; AEO17131; AAS61372.1; -.  
DR PIR; A10315; A10315.  
DR InterPro; IPR007138; ABM.  
DR InterPro; IPR011008; Dimer\_A\_B\_barrel.  
DR Pfam; PF03992; ABM; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 95 AA; 11144 MW; 13192FF15A3A9429 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 95;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
| | | | |  
Db 36 LSYDLFI 42

RESULT 14  
AAS61372 ID AAS61372 PRELIMINARY; PRT; 95 AA.  
AC AAS61372;  
DT 24-MAR-2004 (TrEMBLrel. 27, Created)  
DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN Yp1126.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91001 / Biovar Mediaevalis;  
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,  
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,  
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
RA Yang R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO17131; AAS61372.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 95 AA; 11144 MW; 13192FF15A3A9429 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 95;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
| | | | |  
Db 36 LSYDLFI 42

RESULT 15  
Q6L299 ID Q6L299 PRELIMINARY; PRT; 184 AA.  
AC Q6L299;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocusNames=PT00318;  
OS Picrophilus torridus.  
OC Archaea: Euryarchaeota: Thermoplasmata; Thermoplasmatales;  
OC Picrophilaceae; Picrophilus.  
OX NCBI\_TaxID=82076;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 9790 / ATCC 700027;  
RX PubMed=15184674;  
RA Fuetterer O., Angelov A., Liesegang H., Gottschalk G., Schleper C.,  
RA Schepers B., Dock C., Antranikian G., Liebl W.;  
RT "Genome sequence of Picrophilus torridus and its implications for life  
RT around pH 0.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9091-9096(2004).  
DR EMBL; AE017261; AAT42903.1;-  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 184 AA; 21010 MW; 2332EE9FE3E1A636 CRC64;  
  
Query Match 75.6%; Score 34; DB 2; Length 184;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 1 FLSYDLFV 8  
Db 104 FVAYDLFI 111

Search completed: December 30, 2004, 20:50:55  
Job time : 13.5642 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:34:10 ; Search time 2.06297 Seconds  
(without alignments)  
289.321 Million cell updates/sec

Title: US-10-017-327-11  
Perfect score: 43  
Sequence: 1 NLQLMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	81.4	113	4	US-09-621-976-5972 Sequence 5972, Ap
2	35	81.4	152	4	US-09-513-999C-5230 Sequence 5230, Ap
3	33	76.7	90	4	US-09-248-796A-20334 Sequence 20334, A
4	32	74.4	283	4	US-09-583-110-3146 Sequence 3146, Ap
5	31	72.1	152	4	US-09-583-110-5027 Sequence 5027, Ap
6	31	72.1	159	4	US-09-248-796A-27510 Sequence 27510, A
7	31	72.1	189	4	US-09-248-796A-19441 Sequence 19441, A
8	31	72.1	270	4	US-09-602-787A-676 Sequence 676, App
9	31	72.1	276	4	US-09-252-991A-18006 Sequence 18006, A
10	31	72.1	542	4	US-09-198-452A-496 Sequence 496, App
11	31	72.1	640	4	US-09-248-796A-16129 Sequence 16129, A
12	31	72.1	855	4	US-09-248-796A-18667 Sequence 18667, A
13	30	69.8	67	4	US-09-252-991A-25165 Sequence 25165, A
14	30	69.8	168	4	US-09-270-767-33712 Sequence 33712, A
15	30	69.8	168	4	US-09-270-767-48929 Sequence 48929, A
16	30	69.8	231	4	US-09-540-236-2077 Sequence 2077, Ap
17	30	69.8	412	4	US-09-134-000C-6666 Sequence 6666, Ap
18	30	69.8	656	4	US-09-252-991A-29575 Sequence 29575, A
19	30	69.8	673	3	US-09-134-001C-4480 Sequence 4480, A
20	30	69.8	808	4	US-09-489-039A-11547 Sequence 11547, A
21	30	69.8	1116	4	US-09-583-110-5094 Sequence 5094, Ap
22	30	69.8	1784	3	US-09-040-738-2 Sequence 2, Appli
23	30	69.8	1784	3	US-08-652-426A-2 Sequence 2, Appli
24	29	67.4	133	4	US-09-270-767-38268 Sequence 38268, A
25	29	67.4	133	4	US-09-270-767-53485 Sequence 53485, A
26	29	67.4	228	3	US-09-134-001C-4694 Sequence 4694, Ap
27	29	67.4	228	3	US-09-134-001C-5495 Sequence 5495, Ap

Sequence 93, Appli  
Sequence 30349, A  
Sequence 6, Appli  
Sequence 23698, A  
Sequence 30109, A  
Sequence 15391, A  
Sequence 22687, A  
Sequence 18744, A  
Sequence 29441, A  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 24, Appl  
Sequence 21444, A  
Sequence 3824, Ap  
Sequence 15, Appl  
Sequence 27139, A  
Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-09-621-976-5972  
; Sequence 5972, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621.976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5972  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 75  
; OTHER INFORMATION: Xaa = \* ,Glu,Lys,Gln  
US-09-621-976-5972

Query Match 81.4%; Score 35; DB 4; Length 113;  
Best Local Similarity 87.5%; Pred. No. 5;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLQLMDR 8  
Db 23 NLQLMDR 30

RESULT 2  
US-09-513-999C-5230  
; Sequence 5230, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm

```

; SEQ ID NO 5230
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 80
; OTHER INFORMATION: Xaa=Lys or Gln or Arg
US-09-513-999C-5230

Query Match      81.4%; Score 35; DB 4; Length 152;
Best Local Similarity 87.5%; Pred. No. 7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQLMDRV 8
Db      23 NLQLLVDR 30

RESULT 3
US-09-248-796A-20334
; Sequence 20334, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20334
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20334

Query Match      76.7%; Score 33; DB 4; Length 90;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 NLQLMDRV 9
Db      74 NLQLFMDHI 82

RESULT 4
US-09-583-110-3146
; Sequence 3146, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3146
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3146
```

```

Query Match      74.4%; Score 32; DB 4; Length 283;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 NLQLMDRV 9
Db      172 NLPLLDRI 180

RESULT 5
US-09-583-110-5027
; Sequence 5027, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5027
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5027

Query Match      72.1%; Score 31; DB 4; Length 152;
Best Local Similarity 55.6%; Pred. No. 47;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 NLQLMDRV 9
Db      132 NLQILLRI 140

RESULT 6
US-09-248-796A-27510
; Sequence 27510, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27510
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (151)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-27510

Query Match      72.1%; Score 31; DB 4; Length 159;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 NLQLMDRV 9
|||:|:|
Db 19 NLQILKQV 27

RESULT 7
US-09-248-796A-19441
; Sequence 19441, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.112
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19441
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19441

Query Match 72.1%; Score 31; DB 4; Length 189;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQLMDRV 9
|||:|:|
Db 28 NLQILKQV 36

RESULT 8
US-09-602-787A-676
; Sequence 676, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kr"ger, Burkhard
; APPLICANT: Sch"der, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6

; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 676
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-676

Query Match 72.1%; Score 31; DB 4; Length 270;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQLMDRV 9
|||:|:|
Db 22 NLPLIVDRV 30

RESULT 9
US-09-252-991A-18006
; Sequence 18006, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 18006
/ LENGTH: 276
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: (53)
/ OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-18006

Query Match          72.1%; Score 31; DB 4; Length 276;
Best Local Similarity 66.7%; Pred. No. 91;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NLQLMDRV 9
Db      60 HLELLVDRV 68

RESULT 10
US-09-198-452A-496
/ Sequence 496, Application US/09198452A
/ Patent No. 6559294
/ GENERAL INFORMATION:
/ APPLICANT: Griffois, R.
/ TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
/ TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
/ TITLE OF INVENTION: and treatment of infection
/ FILE REFERENCE: 9710-003-999
/ CURRENT APPLICATION NUMBER: US/09/198,452A
/ CURRENT FILING DATE: 1998-11-24
/ NUMBER OF SEQ ID NOS: 6849
/ SEQ ID NO 496
/ LENGTH: 542.
/ TYPE: PRT
/ ORGANISM: Chlamydia pneumoniae
US-09-198-452A-496

Query Match          72.1%; Score 31; DB 4; Length 542;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 NLQLMDRV 9
Db      531 NLQLASDRI 539

RESULT 11
US-09-248-796A-16129
/ Sequence 16129, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 16129
/ LENGTH: 640
```

```
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-16129

Query Match          72.1%; Score 31; DB 4; Length 640;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NLQLMDRV 9
Db      12 NLKLLIDKI 20

RESULT 12
US-09-248-796A-18667
/ Sequence 18667, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 18667
/ LENGTH: 855
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-18667

Query Match          72.1%; Score 31; DB 4; Length 855;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 NLQLMDRV 9
Db      493 NLRLLLDHV 501

RESULT 13
US-09-252-991A-25165
/ Sequence 25165, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 25165
/ LENGTH: 67
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25165

Query Match          69.8%; Score 30; DB 4; Length 67;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      3 QLLMDRV 9
Db      47 QLVMDRV 53
```

RESULT 14  
US-09-270-767-33712  
; Sequence 33712, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 33712  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-33712

Query Match 69.8%; Score 30; DB 4; Length 168;  
Best Local Similarity 62.5%; Pred. No. 85;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LOLLMDRV 9  
|:|:|:  
Db 84 LELLLDRI 91

RESULT 15  
US-09-270-767-48929  
; Sequence 48929, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48929  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-48929

Query Match 69.8%; Score 30; DB 4; Length 168;  
Best Local Similarity 62.5%; Pred. No. 85;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LOLLMDRV 9  
|:|:|:  
Db 84 LELLLDRI 91

Search completed: December 30, 2004, 20:53:57  
Job time : 2.06297 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:35:11 ; Search time 8.81864 Seconds  
(without alignments)  
367.126 Million cell updates/sec

Title: US-10-017-327-11

Perfect score: 43

Sequence: 1 NLQLMDRV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	9	11 US-09-870-216C-11	Sequence 11, Appl
2	43	100.0	9	13 US-10-017-327-11	Sequence 11, Appl
3	43	100.0	256	9 US-09-925-300-1353	Sequence 1353, Ap
4	43	100.0	352	11 US-09-870-216C-2	Sequence 2, Appl
5	43	100.0	352	13 US-10-017-327-2	Sequence 2, Appl
6	43	100.0	352	16 US-10-408-765A-347	Sequence 347, App
7	36	83.7	311	17 US-10-739-930-6423	Sequence 6423, Ap
8	35	81.4	180	15 US-10-264-237-1883	Sequence 1883, Ap
9	35	81.4	239	9 US-09-925-300-1229	Sequence 1229, Ap
10	34	79.1	371	17 US-10-425-115-239256	Sequence 239256,
11	34	79.1	372	15 US-10-425-114-59828	Sequence 59828, A
12	33	76.7	874	14 US-10-163-214-13	Sequence 13, Appl
13	33	76.7	987	16 US-10-437-963-113606	Sequence 113606,

14	74.4	125	9	US-09-867-550-422	Sequence 422, App
15	74.4	155	17	US-10-425-115-357844	Sequence 357844,
16	74.4	178	15	US-10-424-599-233102	Sequence 233102,
17	74.4	203	14	US-10-369-493-11994	Sequence 11994, A
18	74.4	423	16	US-10-437-963-166526	Sequence 166526,
19	74.4	2697	9	US-09-961-527A-5	Sequence 5, Appl
20	74.4	4080	15	US-10-307-817-138	Sequence 138, App
21	74.4	4624	16	US-10-408-765A-2991	Sequence 2991, Ap
22	72.1	156	16	US-10-437-963-113390	Sequence 113390,
23	72.1	160	17	US-10-425-115-325709	Sequence 325709,
24	72.1	162	16	US-10-437-963-142657	Sequence 142657,
25	72.1	175	16	US-10-437-963-125323	Sequence 125323,
26	72.1	229	17	US-10-425-115-218767	Sequence 218767,
27	72.1	270	9	US-09-738-626-5133	Sequence 5133, Ap
28	72.1	270	15	US-10-627-476-676	Sequence 676, App
29	72.1	320	15	US-10-425-114-63198	Sequence 63198, A
30	72.1	336	17	US-10-425-115-345058	Sequence 345058,
31	72.1	455	17	US-10-739-930-10942	Sequence 10942, A
32	72.1	529	17	US-10-848-111-4	Sequence 4, Appl
33	72.1	542	15	US-10-289-762-496	Sequence 496, App
34	72.1	823	17	US-10-425-115-340748	Sequence 340748,
35	72.1	833	14	US-10-369-493-14960	Sequence 14960, A
36	72.1	838	14	US-10-369-493-11738	Sequence 11738, A
37	72.1	838	14	US-10-369-493-14794	Sequence 14794, A
38	72.1	916	15	US-10-287-226-330	Sequence 330, App
39	72.1	1504	14	US-10-369-493-22466	Sequence 22466, A
40	69.8	55	15	US-10-335-977-6139	Sequence 6139, Ap
41	69.8	130	17	US-10-425-115-362529	Sequence 362529,
42	69.8	132	14	US-10-104-047-2471	Sequence 2471, Ap
43	69.8	161	16	US-10-767-701-50312	Sequence 50312, A
44	69.8	167	14	US-10-183-687-390	Sequence 390, App
45	69.8	181	17	US-10-425-115-331931	Sequence 331931,

#### ALIGNMENTS

RESULT 1  
US-09-870-216C-11  
; Sequence 11, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870.216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-216C-11

Query Match 100.0%; Score 43; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLQLMDRV 9  
| | | | | | | | | |  
Db 1 NLQLMDRV 9

RESULT 2  
US-10-017-327-11  
; Sequence 11, Application US/10017327

[illegible]



```
Query Match      100.0%; Score 43; DB 16; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQLQMDRV 9
Db 242 NQLQMDRV 250
|||||

RESULT 7
US-10-739-930-6423
; Sequence 6423, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USRS THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 6423
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C3789_1.p
US-10-739-930-6423

Query Match      83.7%; Score 36; DB 17; Length 311;
Best Local Similarity 87.5%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQLMDRV 9
Db 146 LQLMDRV 153
|||||

RESULT 8
US-10-264-237-1883
; Sequence 1883, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1883
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-1883

Query Match      81.4%; Score 35; DB 15; Length 180;
Best Local Similarity 87.5%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQLQMDRV 8
Db 23 NQLQMDRV 30
|||||

RESULT 9
US-09-925-300-1229
; Sequence 1229, Application US/09925300
```

```
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1229
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1229

Query Match      81.4%; Score 35; DB 9; Length 239;
Best Local Similarity 87.5%; Pred. No. 74;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQLQMDRV 8
Db 82 NQLQMDRV 89
|||||

RESULT 10
US-10-425-115-239256
; Sequence 239256, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 239256
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_14978C.1.pep
US-10-425-115-239256

Query Match      79.1%; Score 34; DB 17; Length 371;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NQLQMDRV 9
Db 356 NQLQMDRV 364
|||||

RESULT 11
US-10-425-114-59828
; Sequence 59828, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59828
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3911-002-C10_F11.pep
; US-10-425-114-59828

```

```

Query Match      79.1%; Score 34; DB 15; Length 372;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 NLQLMDRV 9
| :|||||
Db 357 NAKLLMDRV 365

```

## RESULT 12

```

US-10-163-214-13
; Sequence 13, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Bucier, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Vigna unguiculata
; US-10-163-214-13

```

```

Query Match      76.7%; Score 33; DB 14; Length 874;
Best Local Similarity 75.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 NLQLMDRV 8
| :|||||
Db 195 NLQLLLDK 202

```

## RESULT 13

```

US-10-437-963-113606
; Sequence 113606, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

```

```

; SEQ ID NO 113606
; LENGTH: 987
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17378C.1.pep
; US-10-437-963-113606

```

```

Query Match      76.7%; Score 33; DB 16; Length 987;
Best Local Similarity 66.7%; Pred. No. 8.2e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 NLQLMDRV 9
| :|||||
Db 564 DLQLLVDR 572

```

## RESULT 14

```

US-09-867-550-422
; Sequence 422, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and f
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-867-550-422

```

```

Query Match      74.4%; Score 32; DB 9; Length 125;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 NLQLMDRV 8
| :|||||
Db 22 NLQLLLDQ 29

```

## RESULT 15

```

US-10-425-115-357844
; Sequence 357844, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 357844
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(155)

```

```
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_89523C.1.pep
US-10-425-115-357844

Query Match      74.4%; Score 32; DB 17; Length 155;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 NLQLMDRV 9
      |||||
Db      144 NLTFLMDRI 152

Search completed: December 30, 2004, 21:00:36
Job time : 10.8186 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:25:59 ; Search time 10.4962 Seconds  
(without alignments)  
307.593 Million cell updates/sec

Title: US-10-017-327-11

Perfect score: 43

Sequence: 1 NLQLMDRV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	9	5	ABB08365	Abb08365 The natur
2	43	100.0	9	7	ABR82217	ABR82217 Human ant
3	43	100.0	256	3	AAB56775	AAB56775 Human pro
4	43	100.0	302	7	ADE29060	ADE29060 Human EIF
5	43	100.0	347	7	ADE29059	ADE29059 Human EIF
6	43	100.0	352	5	ABB08360	Abb08360 Human eIF
7	43	100.0	352	7	ABR82212	ABR82212 Human ant
8	43	100.0	352	7	ADJ68541	ADJ68541 Human hea
9	43	100.0	368	7	ADC31312	ADC31312 Human nov
10	36	83.7	260	3	AAG57646	AAG57646 Arabidops
11	36	83.7	311	3	AGS57645	AGS57645 Arabidops
12	36	83.7	334	3	AAG57644	AAG57644 Arabidops
13	35	81.4	133	4	AAB92665	AAB92665 Human pro
14	35	81.4	152	3	AAG01149	AAG01149 Human sec
15	35	81.4	180	2	AAY28922	AAY28922 Human reg
16	35	81.4	180	5	ABB89507	Abb89507 Human pol
17	35	81.4	180	7	ADD24442	ADD24442 HLA-B*46 I
18	35	81.4	180	7	ADI15876	ADI15876 Human PP
19	35	81.4	180	7	ADI15933	ADI15933 Human PP
20	35	81.4	180	8	ADJ81699	ADJ81699 Tumour an
21	35	81.4	239	3	AAB56651	AAB56651 Human pro
22	35	81.4	1809	7	ADE55015	ADE55015 Rat Prote
23	35	81.4	1809	7	ADE55009	ADE55009 Rat Prote
24	35	81.4	1809	7	ADE55006	ADE55006 Rat Prote
25	35	81.4	1809	7	ADE55012	ADE55012 Rat Prote

26	34	79.1	1116	4	ABB59885	Abb59885 Drosophil
27	33	76.7	509	7	ADM26506	Adm26506 Hyperther
28	33	76.7	874	7	AAE39967	Aae39967 Vigna ung
29	32	74.4	125	5	ABP63841	Abp63841 Human ORF
30	32	74.4	228	4	AAU44957	Aau44957 Propionib
31	32	74.4	228	6	ABM41476	Abm41476 Propionib
32	32	74.4	269	5	ABP28324	Abp28324 Streptoco
33	32	74.4	283	8	ADK46631	Adk46631 Streptoco
34	32	74.4	621	5	ABP28470	Abp28470 Streptoco
35	32	74.4	2697	6	AAO26550	Aao26550 Cochliobo
36	32	74.4	4080	8	ADO42289	Ado42289 Human NOV
37	32	74.4	4624	7	ADE78572	Ade78572 Dynein ax
38	32	74.4	4624	7	ADJ71185	Adj71185 Human hea
39	31	72.1	152	8	ADK48512	Adk48512 Streptoco
40	31	72.1	270	4	AAB76847	Aab76847 Corynebac
41	31	72.1	270	4	AAG91379	Aag91379 C glutami
42	31	72.1	276	7	ABO69260	Abo69260 Pseudomon
43	31	72.1	309	3	AAI57447	Aay57447 Mouse hom
44	31	72.1	403	8	ADM80057	Adm80057 Spiramyci
45	31	72.1	403	8	ADN97573	Adn97573 S ambofac

ALIGNMENTS

RESULT 1  
ABB08365  
ID ABB08365 standard; protein; 9 AA.

AC ABB08365;

DT 07-MAY-2002 (first entry)

DE The natural epitope of human cancer antigen eIF3.

XX Human; melanoma antigen eukaryotic initiation factor 3; eIF3;

KW ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;

KW anti-cancer; vaccine.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1 /note= "HLA-2 binding residue"

FT Domain 2 /note= "HLA-2 binding residue"

FT Domain 3..8 /note= "T-cell receptor (TCR) binding domain"

FT Domain 9 /note= "HLA-2 binding residue"

WO200192307-A2.

06-DEC-2001.

30-MAY-2001; 2001WO-US017456.

31-MAY-2000; 2000US-0209391P.

17-AUG-2000; 2000US-0226258P.

20-DEC-2000; 2000US-0257008P.

(GENZ ) GENZYME CORP.

PI Nicolette CA;

XX WPI; 2002-139606/18.

DR N-PSDB; ABA97216.

XX New therapeutic compounds useful against human ovarian cancer, for

PT modulating immune response in a subject, and for generating antibodies

PT that specifically recognize and bind to these molecules.

XX Claim 10; Page 59; 68pp; English.

PS

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen EIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the natural epitope of human cancer antigen eIF3 (corresponds  
 CC to residues 242-250 of ABB08360)

XX Sequence 9 AA;  
 CC  
 CC Query Match 100.0%; Score 43; DB 5; Length 9;  
 CC Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 9  
 Db 1 NLQLLMDRV 9  
 |||||

RESULT 2  
 ID ABR82217 standard; peptide; 9 AA.  
 AC ABR82217;  
 XX  
 XX 13-OCT-2003 (first entry)  
 XX Human antigen eIF3 epitope (residues 242-250).  
 XX  
 XX Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
 KW cytostatic; gene therapy; human; antigen; epitope.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003050543-A1.  
 FN  
 XX 19-JUN-2003.  
 PD  
 XX  
 XX 05-DEC-2001; 2001WO-US047997.  
 PF  
 XX  
 XX 05-DEC-2001; 2001WO-US047997.  
 PR  
 XX (GENZ ) GENZYME CORP.  
 PA  
 XX  
 XX Nicolette CA;  
 PI  
 XX  
 XX WPI; 2003-532936/50.  
 DR  
 XX  
 XX Aiding in the diagnosis of a neoplastic condition, useful for treating  
 PT cancer and related malignancies comprises determining the amount of  
 PT expression of an eIF3 protein in a test sample isolated from the cell or  
 PT tissue.  
 PT  
 XX  
 XX Claim 6; Page 30; 77pp; English.  
 PS  
 XX  
 XX The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic  
 CC condition or susceptibility to a neoplastic condition based on the amount

CC of expression of the eIF3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. The present sequence  
 CC represents the human antigen eIF3 epitope

XX Sequence 9 AA;  
 CC  
 CC Query Match 100.0%; Score 43; DB 7; Length 9;  
 CC Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 9  
 Db 1 NLQLLMDRV 9  
 |||||

RESULT 3  
 ID AAB56775 standard; protein; 256 AA.  
 XX  
 XX AAB56775;  
 AC  
 XX  
 XX 13-MAR-2001 (first entry)  
 DT  
 XX Human prostate cancer antigen protein sequence SEQ ID NO:1353.  
 DE  
 XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200055174-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX  
 XX 08-MAR-2000; 2000WO-US005988.  
 PF  
 XX  
 XX 12-MAR-1999; 99US-0124270P.  
 PR  
 XX (HJWA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 PA  
 XX  
 XX Rosen CA, Ruben SM;  
 PI  
 XX  
 XX WPI; 2000-587513/55.  
 DR N-PSDB; AAF15978.  
 DR  
 XX  
 XX Prostate cancer associated gene sequences, referred to as prostate cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as prostate cancer.  
 PT  
 XX  
 XX Claim 11; Page 1777-1778; 2338pp; English.  
 PS  
 XX  
 XX AAF15566 to AAF16505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytostatic,  
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention

## RESULT 5

```

XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT 242..250
XX FT /note= "natural epitope of human cancer antigen eIF3"
XX FT Binding-site
XX FT 242 /note= "HLA-2 binding residue"
XX FT Binding-site
XX FT 243 /note= "HLA-2 binding residue"
XX FT Binding-site
XX FT 244..249
XX FT /note= "T-cell receptor (TCR) binding domain"
XX FT Binding-site
XX FT 250 /note= "HLA-2 binding residue"
XX FT
XX FN WO200192307-A2.
XX PD 06-DEC-2001.
XX XX
XX XX 30-MAY-2001; 2001WO-US017456.
XX PF 31-MAY-2000; 2000US-0209391P.
XX PR 17-AUG-2000; 2000US-0226258P.
XX PR 20-DEC-2000; 2000US-0257008P.
XX XX
XX PA (GENZ ) GENZYME CORP.
XX XX
XX XX Nicolette CA;
XX XX WPI; 2002-139606/18.
XX DR N-PSDB; ABA97211.
XX XX
XX XX New therapeutic compounds useful against human ovarian cancer, for
XX PT modulating immune response in a subject, and for generating antibodies
XX PT that specifically recognize and bind to these molecules.
XX XX
XX PS Disclosure; Page 63-64; 69pp; English.
XX XX
XX CC The invention relates to novel therapeutic compounds, that are designed
XX CC to enhance binding to MHC molecules and to enhance immunoregulatory
XX CC properties relative to their natural counterparts. The activity of the
XX CC compounds of the invention may be described as cytostatic and
XX CC immunomodulatory. The compounds are useful against human ovarian cancer,
XX CC for modulating immune response in a subject, and for generating
XX CC antibodies that specifically recognize and bind to these molecules.
XX CC Compositions comprising the compounds are useful as components of anti-
XX CC cancer vaccines and to expand immune effector cells that are specific for
XX CC eukaryotic initiation factor). The peptides or polypeptides conjugated to
XX CC a detectable agent may be used in diagnostic procedures, such as in the
XX CC detection and purification of antibodies, and as immunogens for
XX CC production of antibodies. The polynucleotides can be used as primers for
XX CC detecting genes or gene transcripts expressed in APC to confirm
XX CC transduction of the polynucleotides into host cells. The current sequence
XX CC represents the human melanoma antigen eukaryotic initiation factor 3
XX CC (eIF3) amino acid sequence
XX SQ
XX Sequence 352 AA;
XX Query Match 100.0%; Score 43; DB 5; Length 352;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NLQLLMDRV 9
XX Db 242 NLQLLMDRV 250
XX
XX RESULT 7
XX ABR82212
XX ID ABR82212 standard; protein; 352 AA.
XX XX
XX AC ABR82212;

```

```

XX 13-OCT-2003 (first entry)
XX DT
XX DE Human antigen eukaryotic translation initiation factor 3 (eIF3).
XX KW Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;
XX KW cytostatic; gene therapy; human; antigen.
XX XX
XX OS Homo sapiens.
XX XX WO2003050543-A1.
XX PN
XX XX 19-JUN-2003.
XX PD
XX PF 05-DEC-2001; 2001WO-US047997.
XX XX
XX PR 05-DEC-2001; 2001WO-US047997.
XX XX
XX PA (GENZ ) GENZYME CORP.
XX XX
XX XX Nicolette CA;
XX XX WPI; 2003-532936/50.
XX DR N-PSDB; ACC85029.
XX XX
XX PT Aiding in the diagnosis of a neoplastic condition, useful for treating
XX PT cancer and related malignancies comprises determining the amount of
XX PT expression of an eIF3 protein in a test sample isolated from the cell or
XX PT tissue.
XX XX
XX PS Claim 6; Page 71-72; 77pp; English.
XX XX
XX CC The invention relates to aiding in the diagnosis of a neoplastic
XX CC condition or susceptibility to a neoplastic condition of an animal cell
XX CC or tissue. The method involves determining the amount of expression of an
XX CC eukaryotic translation initiation factor 3 (eIF3) protein in a test
XX CC sample isolated from the cell or tissue, and diagnosing a neoplastic
XX CC condition or susceptibility to a neoplastic condition based on the amount
XX CC of expression of the eIF3 protein. The methods, compounds and kits are
XX CC useful in therapeutics, diagnostic and screening methods for human cancer
XX CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,
XX CC pancreatic or gastrointestinal cancer, or melanoma. The present sequence
XX CC represents the human antigen eIF3
XX XX
XX SQ Sequence 352 AA;
XX Query Match 100.0%; Score 43; DB 7; Length 352;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NLQLLMDRV 9
XX Db 242 NLQLLMDRV 250
XX
XX RESULT 8
XX ADJ68541
XX ID ADJ68541 standard; protein; 352 AA.
XX XX
XX AC ADJ68541;
XX XX
XX DT 06-MAY-2004 (first entry)
XX XX
XX DE Human heat mitochondrial protein as a therapeutic target SeqID347.
XX XX
XX KW mitochondrial; human; screening assay; diabetes mellitus;
XX KW Huntington's disease; osteoarthritis;
XX KW Leber's hereditary optic neuropathy; LHON;
XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX KW myoclonic epilepsy tagged red fibre syndrome; MERRF; cancer;
XX KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
XX KW osteopathic; ophthalmological; cytostatic.
XX XX

```



```

OS Homo sapiens.
XX WO2003087768-A2.
XX
XX 23-OCT-2003.
XX
XX 04-APR-2003; 2003WO-US010870.
XX
XX 12-APR-2002; 2002US-0372843P.
XX
XX 17-JUN-2002; 2002US-0389987P.
XX
XX 20-SEP-2002; 2002US-0412418P.
XX
XX (MITO-) MITOKOR.
XX
XX (BUCK-) BUCK INST AGE RES.
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
XX Warnock DB;
XX
XX MPI; 2003-845369/78.
XX
XX Identifying a mitochondrial target for drug screening assays and for
XX treating diseases associated with altered mitochondrial function.
XX comprises detecting a modified polypeptide in a sample and correlating
XX with the disease.
XX
XX Claim 1; SEQ ID NO 347; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
XX for therapeutic intervention in treating a disease associated with
XX altered mitochondrial function. Specifically, it refers to a method for
XX identifying proteins of the human heart mitochondrial proteome that are
XX useful for drug screening assays, as well as therapeutic targets. The
XX present invention describes a method for identifying such proteins that
XX can be used in the treatment of various diseases associated with altered
XX mitochondrial function including diabetes mellitus, Huntington's disease,
XX osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX compositions have neuroprotective, neurotropic, antidiabetic,
XX anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX cytosstatic activities. This polypeptide sequence is a human heart
XX mitochondrial protein of the invention.
XX
XX Sequence 352 AA;
XX
XX Query Match 100.0%; Score 43; DB 7; Length 352;
XX Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Gaps 0;
XX Matches 9; Conservative 0; Indels 0;
XX
XX Qy 1 NLQLMDRV 9
XX |||||
XX Db 242 NLQLMDRV 250
XX
XX RESULT 9
XX ADC31312
XX ID ADC31312 standard; protein; 368 AA.
XX
XX AC ADC31312;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human novel polypeptide sequence, SEQ ID NO:1394.
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX ulcers; osteoporosis; autoimmune disease; cancer;
XX molecular weight marker; food supplement; antiparkinsonian; neurotropic;
XX neuroprotective; antianemia; anticoagulant; thrombolytic; vulnerary;
XX antifulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX gene therapy; chromosome 8p22.
XX

```

---

```

OS Homo sapiens.
XX WO2003029271-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030474.
XX
XX 24-SEP-2001; 2001US-0324631P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX
XX MPI; 2003-371981/35.
XX
XX N-PSDB; ADC30341.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
XX treating conditions such as neurodegenerative diseases, anemias, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.
XX
XX Claim 20; SEQ ID NO 1394; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 99% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX contig sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating diseases such as Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,
XX and as food supplements. The present sequence represents a specifically
XX claimed human polypeptide sequence of the invention. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 368 AA;
XX
XX Query Match 100.0%; Score 43; DB 7; Length 368;
XX Best Local Similarity 100.0%; Pred. No. 2;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 NLQLMDRV 9
XX |||||
XX Db 258 NLQLMDRV 266
XX
XX RESULT 10
XX AAG57646
XX ID AAG57646 standard; protein; 260 AA.
XX

```





PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140685P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145152P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145819P.  
PR 28-JUL-1999; 99US-0145851P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 08-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149802P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 83.7%; Score 36; DB 3; Length 311;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLMDRV 9  
DB 146 LQLMDRI 153

RESULT 12

AGS57644  
ID AAGS57644 standard; protein; 334 AA.

XX AAGS57644;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 74310.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-00301439.  
XX  
XX  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-01231180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 23-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 18-APR-1999; 99US-01287114P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145193P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.

```

PR 22-SEP-1999; 99US-01551139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159594P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      83.7%; Score 36; DB 3; Length 334;
Best Local Similarity 87.5%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLMDRV 9
DB 169 LQLMDRI 176
|||||:

RESULT 13
AAB92665
ID AAB92665 standard; protein; 133 AA.
AC AAB92665;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:11024.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.
PR (HELI-) HELIX RES INST.
PA
XX
PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
PT
XX
PS Claim 8; SEQ ID NO 11024; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 133 AA;

Query Match      81.4%; Score 35; DB 4; Length 133;
Best Local Similarity 87.5%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLMDR 8
DB 23 NLQLLVDR 30
|||||:

RESULT 14
AAG01149
ID AAG01149 standard; protein; 152 AA.
XX
AC AAG01149;
XX
DT 06-OCT-2000 (first entry)
DE Human secreted protein, SEQ ID NO: 5230.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX

```

```

PA (GSEST ) GENSET.
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX N-PSDB; AAC01155.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 13; SEQ ID NO 5230; 71pp + Sequence Listing; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number of
XX 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
XX prepared from total human RNAs or polyA+ RNAs derived from 30 different
XX tissues. EST sequences usually correspond mainly to the 3' untranslated
XX region (UTR) of the mRNA because they are often obtained from oligo-dT
XX primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
XX sequences derived from the 5' ends of mRNAs and even in those cases where
XX longer cDNA sequences have been obtained, the full 5' UTR is rarely
XX included. 5' ESTs are derived from mRNAs with intact 5' ends and can
XX therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
XX are also used in diagnostic, forensic, gene therapy and chromosome
XX mapping procedures. They are used to obtain upstream regulatory sequences
XX and to design expression and secretion vectors
XX
XX Sequence 152 AA;
XX
XX Query Match 81.4%; Score 35; DB 3; Length 152;
XX Best Local Similarity 87.5%; Pred. NO. 34;
XX Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 NLQLLMDR 8
XX |||||:
XX 23 NLQLLVDR 30
XX
XX RESULT 15
XX AAY28922
XX ID AAY28922 standard; protein; 180 AA.
XX
XX AC AAY28922;
XX
XX DT 21-SEP-1999 (first entry)
XX
XX DE Human regulatory protein HRGP-8.
XX
XX KW Regulatory protein; HRGP; human; cell proliferation; immune response;
XX cancer; adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;
XX tetracarcinoma; Addison's disease; adult respiratory distress syndrome;
XX allergy; anemia; asthma; atherosclerosis; bronchitis; cholecystitis;
XX Crohn's disease; ulcerative colitis; atopic dermatitis; dermatomyositis;
XX diabetes mellitus; emphysema; atrophic gastritis; glomerulonephritis;
XX gout; Grave's disease; hyperinophillia; irritable bowel syndrome; AIDS;
XX lupus erythematosus; multiple sclerosis; myasthenia gravis; infarction;
XX osteoarthritis; osteoporosis; pancreatitis; polymyositis; hemodialysis;
XX arthritis; scleroderma; Sjogren's syndrome; and autoimmune thyroiditis;
XX infection; trauma.
XX
XX OS Homo sapiens.
XX
XX PN W09933870-A2.
XX
XX PD 08-JUL-1999.
XX
XX PF 22-DEC-1998; 98WO-US027471.
XX
XX PR 31-DEC-1997; 97US-00001403.
XX
XX PA (INCY-) INCYTE PHARM INC.
XX
XX PI Lal P, Bandman O, Hillman JL, Au-Young J, Tang YT, Yue H;

```

---

```

PI Shah P, Guegler KJ, Corley NC;
XX
XX WPI; 1999-430229/36.
XX N-PSDB; AAX89292.
XX
XX New human regulatory proteins, useful for diagnosing, preventing and
XX treating disorders associated with expression of regulatory proteins.
XX
XX Claim 1; Page 77; 89pp; English.
XX
XX The invention provides novel human regulatory proteins (AAY28915-926),
XX designated HRGP, and their polynucleotides (AAX89285-296). The proteins of
XX HRGP is closely associated with cell proliferation and the polypeptides
XX and polynucleotides are useful in the diagnosis, treatment and prevention
XX of diseases associated with cell proliferation, particularly immune
XX responses and cancer. The protein or agonists may be administered to
XX treat or prevent a cancer such as adenocarcinoma, leukemia, lymphoma, but
XX melanoma, myeloma, sarcoma, and tetracarcinoma. Such cancers include, but
XX are not limited to, cancers of the adrenal gland, bladder, bone, brain,
XX cervix, breast, gall bladder, ganglia, ovary and pancreas. Where HRGP is
XX promoting leukocyte activity or proliferation, antagonists which decrease
XX the activity of HRGP are administered. Such responses may be associated
XX with disorders such as Addison's disease, adult respiratory distress
XX syndrome, AIDS, allergies, anemia, asthma, atherosclerosis, bronchitis,
XX cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis,
XX dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, gout,
XX glomerulonephritis, Grave's disease, hyperinophillia, irritable bowel
XX syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis,
XX myocardial or pericardial infarction, osteoarthritis, osteoporosis,
XX pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjogren's
XX syndrome, and autoimmune thyroiditis, complications of cancer,
XX hemodialysis, extracorporeal circulation; viral, bacterial, fungal,
XX parasitic, protozoan, and helminthic infections; and trauma. The HRGP
XX polynucleotide may be used for diagnosis of these conditions, and as a
XX source of primers and probes
XX
XX SQ Sequence 180 AA;
XX
XX Query Match 81.4%; Score 35; DB 2; Length 180;
XX Best Local Similarity 87.5%; Pred. No. 41;
XX Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 NLQLLMDR 8
XX |||||:
XX 23 NLQLLVDR 30
XX
XX Search completed: December 30, 2004, 20:43:07
XX Job time : 13.4962 secs

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:30:54 ; Search time 1.97229 Seconds  
(without alignments)  
439.058 Million cell updates/sec

Title: US-10-017-327-11

Perfect score: 43

Sequence: 1 NLQLMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	83.7	311	2 G86324	hypothetical prote
2	35	81.4	1809	2 S57329	tuberosus sclerosis
3	34	79.1	1116	2 T13854	nuclear protein SA
4	32	74.4	283	2 E98065	conserved hypotet
5	32	74.4	305	2 F69306	hypothetical prote
6	32	74.4	572	2 S73730	MG307 homolog H08
7	32	74.4	1244	2 S73731	probable lipoprote
8	31	72.1	258	2 B72697	hypothetical prote
9	31	72.1	270	2 D69791	conserved hypotet
10	31	72.1	291	2 AE2416	hypothetical prote
11	31	72.1	445	1 S74826	NADH2 dehydrogenas
12	31	72.1	467	2 C84420	hypothetical prote
13	31	72.1	519	2 E84811	hypothetical prote
14	31	72.1	582	2 A50192	conserved hypotet
15	31	72.1	651	2 JN0657	methyltransferase
16	31	72.1	652	2 AD0546	site-specific DNA-
17	31	72.1	672	2 A72076	hypothetical prote
18	31	72.1	672	2 B86548	hypothetical prote
19	31	72.1	841	2 JQ0647	preprotein translo
20	31	72.1	861	2 E57473	1708 [imported] -
21	31	72.1	861	2 A12691	copper transportin
22	31	72.1	1504	2 A33602	DNA-directed DNA p
23	30	69.8	129	2 T09329	XilF1 protein - hu
24	30	69.8	144	2 F91130	hypothetical prote
25	30	69.8	144	2 F85975	hypothetical prote
26	30	69.8	217	2 T44404	adenylate kinase (
27	30	69.8	218	2 AE1147	carboxylesterase h
28	30	69.8	219	2 AE1506	weakly carboxylest
29	30	69.8	227	2 A75304	ABC transporter, A

ALIGNMENTS

RESULT 1

G86324 hypothetical protein T29M8.1 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: G86324

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86324

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <STO>

A:Cross-references: UNIPROT:Q9LMB2; GB:AE005172; NID:g8954052; PIDN:AAF82225.1; GSPDB:GN

C:Genetics:

A:Map position: 1

Query Match 83.7%; Score 36; DB 2; Length 311;

Best Local Similarity 87.5%; Pred. No. 4.8;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQLMDRV 9

Db 146 LQLMDRI 153

RESULT 2

S57329

tuberosus sclerosis 2 homolog - rat

N:Alternate names: tuberin

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004

C:Accession: S57329; I52834

R:Kobayashi, T.; Nishizawa, M.; Hirayama, Y.; Kobayashi, E.; Hino, O.

Nucleic Acids Res. 23, 2608-2613, 1995

A:Title: cDNA structure, alternative splicing and exon-intron organization of the predi

A:Reference number: S57329; MUID:95380273; PMID:7651821

A:Accession: S57329

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1809 <KOB>

A:Cross-references: UNIPROT:P49816; EMBL:D50413; NID:gl181080; PIDN:BAA08914.1; PTD:g994

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

A>Note: only a part of the coding sequence is given



A;Accession: S73731  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1244 <HIM>  
A;Cross-references: UNIPROT:P75342; EMBL:AE000039; GB:U00089; NID:gi1674082; PIDN:AA9605  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C;Genetics:  
A;Genetic code: SGC3  
C;Superfamily: Mycoplasma hypothetical protein MG309

Query Match 74.4%; Score 32; DB 2; Length 1244;  
Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQLMDRV 9  
| | | | | | | | | |  
Db 1082 NNQLLDRI 1090

RESULT 8  
B72697  
hypothetical protein APE0994 - Aeropyrum pernix (strain KI)  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: B72697  
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: B72697  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-258 <KAW>  
A;Cross-references: UNIPROT:Q9YDB9; DDBJ:AP000060; NID:gs104188; PIDN:BA799978.1; PID:di  
A;Experimental source: strain KI  
C;Genetics:  
A;Gene: APR0994

Query Match 72.1%; Score 31; DB 2; Length 258;  
Best Local Similarity 75.0%; Pred. No. 48;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQLMDRV 9  
| | | | | | | | | |  
Db 224 LRILMDRV 231

RESULT 9  
D69791  
conserved hypothetical protein yead - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: D69791  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C;Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsreim, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Sero  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: D69791  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA

A;Residues: 1-270 <KUN>  
A;Cross-references: UNIPROT:P94475; GB:Z39107; GB:AL009126; NID:g2632866; PIDN:CAB12453.  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yead

Query Match 72.1%; Score 31; DB 2; Length 270;  
Best Local Similarity 87.5%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQLMDR 8  
| | | | | | | |  
Db 115 NLFLMDR 122

RESULT 10  
AE2416  
hypothetical protein all4885 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AE2416  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AE2416  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-291 <KUR>  
A;Cross-references: UNIPROT:Q8YMP8; GB:BA000019; PIDN:BAB76584.1; PID:gi17134022; GSPDB:G  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all4885

Query Match 72.1%; Score 31; DB 2; Length 291;  
Best Local Similarity 75.0%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQLMDR 8  
| | | | | | | |  
Db 170 NLQKIMDR 177

RESULT 11  
S74826  
NADH2 dehydrogenase (EC 1.6.99.3) - Synechocystis sp. (strain PCC 6803)  
N;Alternate names: protein slr0851  
A;Species: Synechocystis sp.  
A;Variety: PCC 6803  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S74826  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
s.  
A;Reference number: S74322; MUID:97061201; PMID:8905231  
A;Accession: S74826  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-445 <KAN>  
A;Cross-references: UNIPROT:P73739; EMBL:D90909; GB:AB001339; NID:gi652844; PIDN:BAAL1778  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C;Genetics:  
A;Gene: ndh  
C;Superfamily: NADH dehydrogenase  
C;Keywords: electron transfer; FAD; flavoprotein; NAD; oxidoreductase

Query Match 72.1%; Score 31; DB 1; Length 445;  
Best Local Similarity 55.6%; Pred. No. 88;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 9  
|.:|.:|.:|  
Db 76 NLRVLMDRV 84

## RESULT 12

C84420  
hypothetical protein At2g01070 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: C84420  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84420  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-467 <STO>  
A:Cross-references: UNIPROT:Q9SVJ9; GB:AE002093; NID:G6598622; PIDN:AAF18655.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g01070  
A:Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 467;  
Best Local Similarity 62.5%; Pred. No. 93;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 8  
|.:|.:|.:|  
Db 411 NWKLMMDR 418

## RESULT 13

E84811  
hypothetical protein At2g38960 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: E84811  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: E84811  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-519 <STO>  
A:Cross-references: UNIPROT:Q92V11; GB:AE002093; NID:G3928083; PIDN:AAC79609.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g38960  
A:Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 519;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 9  
|.:|.:|.:|  
Db 267 NLTLMDRV 275

## RESULT 14

A90192  
conserved hypothetical protein [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: A90192  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P. arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: A90192  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-582 <KUR>  
A:Cross-references: UNIPROT:Q97ZV6; GB:AE006641; NID:gl3813624; PIDN:AAK40792.1; GSPDB:GN  
C:Genetics:  
A:Gene: SS00469

Query Match 72.1%; Score 31; DB 2; Length 582;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 9  
|.:|.:|.:|  
Db 571 NLDLMDRI 579

## RESULT 15

JN0657  
methyltransferase (EC 2.1.1.-) - Salmonella typhimurium  
C:Species: Salmonella typhimurium  
C:Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 09-Jul-2004  
C:Accession: JN0657  
R:Dartois, V.; De Backer, O.; Colson, C.  
Gene 127, 105-110, 1993  
A:Title: Sequence of the Salmonella typhimurium StylT1 restriction-modification genes: Hc

A:Reference number: JN0657; MUID:93252265; PMID:8387444  
A:Accession: JN0657  
A:Molecule type: DNA  
A:Residues: 1-651 <DAR>  
A:Cross-references: UNIPROT:P40814; GB:M90544  
C:Comment: This enzyme belongs to the type-III restriction-modification system and is nec

C:Genetics:  
A:Gene: mod

C:Keywords: DNA binding; methyltransferase; restriction modification system  
F:134-137/Region: S-adenosylmethionine binding #status predicted

Query Match 72.1%; Score 31; DB 2; Length 651;  
Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQLLMDRV 9  
|.:|.:|.:|  
Db 215 NLKLMMDRI 223

Search completed: December 30, 2004, 20:52:24  
Job time : 3.97229 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 20:29:59 ; Search time 10.5642 Seconds  
(without alignments)  
490.180 Million cell updates/sec

Title: US-10-017-327-11  
Perfect score: 43  
Sequence: 1 NLQLMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	352	1	IF33 HUMAN	O15372 homo sapien
2	43	100.0	352	2	Q6P9U8	Q6P9U8 rattus norv
3	43	100.0	352	2	AAH60586	AAH60586 rattus no
4	43	100.0	352	2	CAG33187	Cag33187 homo sapi
5	38	88.4	335	2	Q6P381	Q6P381 xenopus tro
6	38	88.4	335	2	AAH64151	AAH64151 xenopus t
7	38	88.4	352	1	IF33 MOUSE	Q91wk2 mus musculu
8	38	88.4	352	2	Q8BTX5	Q8BTX5 mus musculu
9	37	86.0	473	2	Q8DJU7	Q8dju7 synechococc
10	36	83.7	196	2	Q94BU3	Q94bu3 arabidopsis
11	36	83.7	311	2	Q8LAF2	Q8lap2 arabidopsis
12	36	83.7	311	2	Q9LMB2	Q9lmb2 arabidopsis
13	36	83.7	342	2	Q9LMB3	Q9lmb3 arabidopsis
14	35	81.4	133	2	Q9NZ20	Q9nzz0 homo sapien
15	35	81.4	180	2	Q9Y221	Q9y221 homo sapien
16	35	81.4	180	2	Q9WV50	Q9wv50 rattus norv
17	35	81.4	180	2	Q9CXK8	Q9cxk8 mus musculu
18	35	81.4	180	2	Q9D1B4	Q9dlb4 mus musculu
19	35	81.4	180	2	AAH59114	AAH59114 rattus no
20	35	81.4	180	2	BAD05056	BAD05056 homo sapi
21	35	81.4	1742	2	Q7TT21	Q7tt21 mus musculu
22	35	81.4	1742	2	AAH60701	AAH60701 mus muscu
23	35	81.4	1809	1	TSC2 RAT	Q49816 rattus norv
24	35	81.4	1814	1	TSC2 MOUSE	Q61037 mus musculu
25	34	79.1	149	2	Q6Y126	Q6y1z6 pagrus majo
26	34	79.1	149	2	RAP20218	RAP20218 pagrus ma
27	34	79.1	285	2	Q31388	Q31388 cyprinus ca
28	34	79.1	285	2	Q73MF8	Q73mf8 treponema d
29	34	79.1	560	2	AAH12067	AAH12067 treponema
30	34	79.1	681	1	RPOC ANTFO	Q85c16 anthoceros
31	34	79.1	808	2	Q6BZ11	Q6bz11 debaryomyce

32	34	79.1	964	2	Q7UJ58	Q7uj58 rhodopirell
33	34	79.1	1116	2	O18415	O18415 drosophila
34	34	79.1	1127	2	Q9VM62	Q9vm62 drosophila
35	34	79.1	1127	2	AAF52463	Aaf52463 drosophil
36	33	76.7	257	2	Q9MBR8	Q9mbr8 staphylococ
37	33	76.7	345	2	Q6F7M9	Q6f7m9 acinetobact
38	33	76.7	509	2	Q8TWC3	Q8twc3 methanopyru
39	33	76.7	874	2	Q9XGC1	Q9xgc1 vigna ungui
40	32	74.4	35	2	Q88G77	Q88g77 pseudomonas
41	32	74.4	99	2	Q7SEW5	Q7afm5 neurospora
42	32	74.4	99	2	CAB76192	Cae76192 neurospor
43	32	74.4	175	2	Q6ZSS5	Q6zss5 homo sapien
44	32	74.4	175	2	BAC86871	BAC86871 homo sapi
45	32	74.4	181	2	Q6BGJ7	Q6bgj7 paramecium

ALIGNMENTS

RESULT 1  
IF33 HUMAN STANDARD; PRT; 352 AA.  
AC O15372;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DB Eukaryotic translation initiation factor 3 subunit 3 (eIF-3 gamma)  
DE (eIF3 p40 subunit) (eIF3h).  
GN Name=EIF3S3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=98001678; PubMed=9341143;  
RA Asano K., Vornlocher H.-P., Richter-Cook N.J., Merrick W.C.,  
RA Hinnebusch A.G., Hershey J.W.B.;  
RT "Structure of cDNAs encoding human eukaryotic initiation factor 3  
subunits. Possible roles in RNA binding and macromolecular assembly.";  
RL J. Biol. Chem. 272:27042-27052(1997).  
[2]  
RP SEQUENCE FROM N.A.  
RX Schmidt O.G., von Holtum D., Gross S., Horsthemke B., Luedicke H.-J.;  
RT "The gene encoding the p40 subunit of the translation initiation  
factor eIF3 has 8 exons, maps to the Langer-Giedion syndrome region on  
chromosome 9q24, but is not the TRPS gene.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaudo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Gough G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).



```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060586; AAH60586.1; -.
KW Initiation factor.
SQ SEQUENCE 352 AA; 39905 MW; C06307269ADB343 CRC64;

Query Match      100.0%; Score 43; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQLLMDRV 9
Db 242 NQLLMDRV 250

RESULT 4
CAG33187
ID CAG33187 PRELIMINARY; PRT; 352 AA.
AC CAG33187;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE EIF3S3 protein.
GN EIF3S3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
RT vector (pDONR201).";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR456906; CAG33187.1; -.
SQ SEQUENCE 352 AA; 39930 MW; F3A6EFA0CEF587D0 CRC64;

Query Match      100.0%; Score 43; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQLLMDRV 9
Db 242 NQLLMDRV 250

RESULT 5
Q6P381
ID Q6P381 PRELIMINARY; PRT; 335 AA.
AC Q6P381;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75580.
GN Name=MGC75580;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC064151; AAH64151.1; -.
DR InterPro; IPR000555; Mov34_MPN_PAD1.
DR Pfam; PF01398; Mov34; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Hypothetical protein.
SQ SEQUENCE 335 AA; 38538 MW; 18D4D446D6561BE9 CRC64;

Query Match      88.4%; Score 38; DB 2; Length 335;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQLLMDRV 9
Db 225 SLQLMDRV 233

RESULT 6
AAH64151
ID AAH64151 PRELIMINARY; PRT; 335 AA.
AC AAH64151;
DT 25-MAR-2004 (TrEMBLrel. 27, Created)
DT 25-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75580.
GN MGC75580.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```



```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC064151; AAH64151.1; -.
KW Hypothetical protein.
SQ SEQUENCE 335 AA; 38538 MW; 18D4D446D6561BE9 CRC64;

Query Match      88.4%; Score 38; DB 2; Length 335;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLMDRV 9
   :|||||
Db 225 SLQLMDRV 233

RESULT 7
ID IF33 MOUSE STANDARD; PRT; 352 AA.
AC Q91WK2
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 3 (eIF-3 gamma)
DE (eIF3 p40 subunit) (eIF3h).
GN Name=Eif3g3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eyes, and Retina;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Binds to the 40S ribosome and promotes the binding of
methionyl-tRNAi and mRNA. Associates with the p170 subunit of
EIF3.
CC -1- SUBUNIT: eIF-3 is composed of at least 12 different subunits.
CC -1- SIMILARITY: Contains 1 MPN (JAB/Mov34) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC014755; AAH14755.1; -.
DR MEROPS; M67.971; -.
DR MGD; MGI:1915385; Eif3g3.

DR InterPro; IPR003639; Mov34-1.
DR InterPro; IPR000555; Mov34_MPN_PAD1.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD363422; Mov34_1; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Initiation factor; Protein biosynthesis.
SQ SEQUENCE 352 AA; 39832 MW; 96F5AB8E2F41F838 CRC64;

Query Match      88.4%; Score 38; DB 1; Length 352;
Best Local Similarity 88.9%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLMDRV 9
   :|||||
Db 242 SLQLMDRV 250

RESULT 8
Q8BTX5 PRELIMINARY; PRT; 352 AA.
AC Q8BTX5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
length enriched library, clone:E430016K01 product:eukaryotic
DE translation initiation factor 3, subunit 3 (gamma, 40kD), full insert
DE sequence.
GN Name=Eif3g3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA The FANTOM Consortium;
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa E., Watahiki M.,
```





RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY087694; AAMG5231.1; -

KW Hypothetical protein.

SQ SEQUENCE 311 AA; 34387 MW; 1ACCD30D7260A0A0 CRC64;

Query Match 83.7%; Score 36; DB 2; Length 311;

Best Local Similarity 87.5%; Pred. No. 57;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLLMDRV 9

|||||:

Db 146 LQLLMDRI 153

RESULT 12

Q9LMB2

ID Q9LMB2 PRELIMINARY; PRT; 311 AA.

AC Q9LMB2;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE T29M8.1 protein.

GN Name=T29M8.1;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Liu S.X., Vaysberg M., Etgu P., Lee J.M., Lenz C., Pham P., Sakano H.,

RA Toriumi M., Yu G., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,

RA Altafi H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,

RA Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,

RA Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,

RA Federspiel N.A., Theologis A.;

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC069143; AAF82225.1; -

DR PIR; G86324; G86324.

SQ SEQUENCE 311 AA; 34341 MW; F4CCA5BD69623AA0 CRC64;

Query Match

Best Local Similarity 83.7%; Score 36; DB 2; Length 311;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLLMDRV 9

|||||:

Db 146 LQLLMDRI 153

RESULT 13

Q9LMB3

ID Q9LMB3 PRELIMINARY; PRT; 342 AA.

AC Q9LMB3;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE F14D16.30.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,

RA Kim C., Altafi H., Bel Q., Chin C., Chio J., Choi E., Conn L.,

RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,

RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,

RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,

RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,

RA Ecker J.R.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

[2]

RP SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,

RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chio J., Choi E.,

RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,

RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,

RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,

RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,

RA Theologis A., Ecker J.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC068602; AAF79297.1; -

SQ SEQUENCE 342 AA; 38937 MW; 48B2B19ECF26617D CRC64;

Query Match

Best Local Similarity 83.7%; Score 36; DB 2; Length 342;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLLMDRV 9

|||||:

Db 146 LQLLMDRI 153

RESULT 14

Q9NZZ0

ID Q9NZZ0 PRELIMINARY; PRT; 133 AA.

AC Q9NZZ0;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE HSPC180.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20499367; PubMed=11042152;

RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,

RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,

RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.;

RT "Cloning and functional analysis of cDNAs with open reading frames for

RT 300 previously undefined genes expressed in CD34+ hematopoietic

RT stem/progenitor cells.";

RL Genome Res. 10:1546-1560(2000).

DR EMBL; AF161528; AAF29143.1; -

DR SWISS-PROT; Q9NZZ0; -

SQ SEQUENCE 133 AA; 15221 MW; 9C538A2DED755248 CRC64;

Query Match

Best Local Similarity 81.4%; Score 35; DB 2; Length 133;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQLMDR 8

|||||:

Db 23 NLQLMDR 30

RESULT 15

Q9V221

ID Q9V221 PRELIMINARY; PRT; 180 AA.

AC Q9V221;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE HSPC031 (OK/SW-CL.78) (CGI-37 protein) (Hypothetical protein) (OK/SW-

DE cl.76 protein).

GN Name=OK/SW-cl.78; Synonyms=OK/SW-cl.76;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=20499367; PubMed=11042152;  
RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,  
RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,  
RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.,  
RT "Cloning and functional analysis of cDNAs with open reading frames for  
RT 300 previously undefined genes expressed in CD34+ hematopoietic  
RT stem/progenitor cells.";  
RL Genome Res. 10:1546-1560(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Zhang Q., Ye M., Fu G., Zhou J., Guan Z., Huang Q., Xu S., He K.,  
RA Chen S., Shen Y., Chen Z.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Shichijo S., Itoh K.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20272150; PubMed=10810093;  
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;  
RT "Identification of novel human genes evolutionarily conserved in  
RT Caenorhabditis elegans by comparative proteomics.";  
RL Genome Res. 10:703-713(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Myokai F.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF085360; AAD40195.1; -;  
DR EMBL; AB062479; BAB93504.1; -;  
DR EMBL; AF132971; AAD27746.1; -;  
DR EMBL; BC015941; AAH15941.1; -;  
DR EMBL; AB112439; BAD05056.1; -;  
DR EMBL; AB062398; BAB93485.1; -;  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR InterPro; IPR002478; PUA.  
DR InterPro; IPR005155; UPF0113.  
DR Pfam; PF03657; UPF0113; 1.  
DR SMART; SMO0359; PUA; 1.  
DR PROSITE; PS00890; PUA; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 180 AA; 20462 MW; FC6CFB2250AA4FC9 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 180;  
Best Local Similarity 87.5%; Pred. No. 54;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 NLQLMDR 8  
Db 23 NLQLLVDR 30  
|||||:  
|||||:

Search completed: December 30, 2004, 20:50:57  
Job time : 12.5642 secs

**This Page Blank (uspto)**